Applied Neuropsychology: Adult

ISSN: 2327-9095 (Print) 2327-9109 (Online) Journal homepage: www.tandfonline.com/journals/hapn21

Learning disability documentation submitted by osteopathic medical students

Allyson G. Harrison, Benjamin J. Lovett, Shelby Keiser & Irene T. Armstrong

To cite this article: Allyson G. Harrison, Benjamin J. Lovett, Shelby Keiser & Irene T. Armstrong (2021) Learning disability documentation submitted by osteopathic medical students, Applied Neuropsychology: Adult, 28:2, 245-256, DOI: 10.1080/23279095.2019.1627358

To link to this article: https://doi.org/10.1080/23279095.2019.1627358

APPLIED NEUROPSYCHOLOGY: ADULT
2021, VOL. 28, NO. 2, 245–256
https://doi.org/10.1080/23279095.2019.1627358

Routledge
Taylor & Francis Group

Learning disability documentation submitted by osteopathic medical students

Allyson G. Harrison[a], Benjamin J. Lovett[b], Shelby Keiser[c], and Irene T. Armstrong[a]

[a]Regional Assessment & Resource Centre, Queens University, Kingston, Ontario, Canada; [b]Teachers College, Columbia University, New York, NY; [c]Keiser Consulting, Ardmore, Pennsylvania, USA

ABSTRACT

An increasing number of individuals are requesting disability accommodations on high-stakes licensing examinations, and neuropsychological evaluation reports are a frequent component of disability documentation. This study reviewed 103 requests for accommodations based on a learning disability (LD), submitted to the National Board of Osteopathic Medical Examiners between 2013 and 2016 for accommodation on the licensure exams for osteopathic physicians. Depending on criteria employed, anywhere from 84 to 97% of applicants failed to meet recognized criteria for a LD diagnosis, with the vast majority demonstrating academic functioning that was average or better relative to most other individuals in the general population. Thus, although given a disability diagnosis, the majority lacked evidence of substantial impairment that would require accommodations under the Americans with Disabilities Act. Additionally, no significant differences were found between the academic achievement or fluency scores of those requesting 50% as compared to 100% extra time. Finally, many evaluators employed grade-based assessment measures to demonstrate academic impairment, even when candidates' functioning using age-based comparisons was unimpaired. Recommendations for future research and practice based on these and other findings are discussed.

KEYWORDS

Academic accommodations; assessment/diagnosis; cognitive/learning; professional training; validity testing

In order to become licensed as a Doctor of Osteopathic Medicine (D.O.) in the United States, prospective applicants must pass a series of four licensing examinations, known as the Comprehensive Osteopathic Medical Licensing Examination (COMLEX-USA), developed and administered by the National Board of Osteopathic Medical Examiners (NBOME). These exams assess a physician's ability to recall and apply knowledge, concepts, and principles "without hesitation and in a fluid manner," to demonstrate patient-centered skills and competencies fundamental to the practice of osteopathic medicine and that constitute the basis of safe and effective medical practice. Indeed, the mission of NBOME is "to protect the public by providing the means to assess competencies of osteopathic medicine and related health care professions." (National Board of Osteopathic Medical Examiners, 2018, p. 1).

Medical students with disabilities who wish to take the COMLEX exams are legally entitled to appropriate accommodations if their disability conditions substantially limit their ability to access the exams compared to most people in the general population. However, such accommodations should not fundamentally alter the assessment of the essential skills being evaluated (Roberts, 2012). To obtain accommodations as an adult, most testing agencies require a recent diagnostic evaluation that documents a current impairment in a function relevant to the requested accommodation (Lindstrom & Lindstrom, 2017). When used properly, such accommodations should lead to more valid test scores for examinees with disabilities; if given inappropriately, however, they may provide examinees with an unfair advantage, or otherwise compromise the integrity of the examinations (Lovett & Lewandowski, 2015).

To ensure that accommodations are not provided unfairly or inappropriately, accommodations are proffered only to students who meet the legal definition of "disability." U.S. federal laws define a disability as a mental or physical condition that substantially limits one or more major life activities, as compared to the

CONTACT Allyson G. Harrison harrisna@queensu.ca Regional Assessment & Resource Centre, Queens University, Mackintosh-Corry Hall, 68 University Ave, Kingston, Ontario K7L 3N9, Canada.

average person or most people in the general population (Lovett, Gordon, & Lewandowski, 2016). This last component of the definition is especially relevant when evaluating postsecondary students, since only a subset of the general population progresses through postsecondary education.

Learning disabilities and accommodations

Learning disabilities (LDs) are neurodevelopmental disorders that emerge in childhood, interfere with the acquisition and expression of academic skills, and substantially limit educational or occupational functioning (American Psychiatric Association, 2013). Studies show that although students with LDs typically have average or better intellectual ability, they are less likely to complete high school and college, and are more likely to be unemployed or underemployed than are nondisabled individuals (Cortiella, 2013; Cortiella & Horowitz, 2014; Kortering & Braziel, 1999; Rojewski, 1999). Despite these findings, the proportion of postsecondary students reporting LD diagnoses and requesting accommodations at the postsecondary level has increased markedly over the past two decades, and LDs are the most common disability for which academic accommodations are requested in both postsecondary education (Harrison & Wolforth, 2012) and on high-stakes exams (U.S. Government Accountability Office, 2011).The academic accommodation most commonly requested by such students is extended testing time (Harrison & Wolforth, 2012; U.S. Government Accountability Office, 2011).

Some postsecondary students with learning disabilities are not diagnosed until college, raising the question of whether the students' academic skills are actually impaired in an absolute sense, or whether instead their skills are simply insufficient to succeed in more rigorous postsecondary coursework (see, e.g., Sparks & Lovett, 2009b). For other students, whose LD diagnoses were initially made in childhood, any prior accommodations were likely provided under special education laws that are quite different in purpose from disability discrimination laws that protect postsecondary students. Whereas current special education laws were designed to improve the educational achievement of students with disabilities, and these laws judge the appropriateness of an individualized educational program in terms of the student's consequent success (Yell, 2019), disability discrimination laws were instead created to allow individuals with disabilities to compete fairly in employment and other settings, and that fairness is not judged by whether or not individuals with disabilities succeed (Lovett et al., 2016). In addition, the criteria for identifying LD under special education laws vary widely from state to state (Maki, Floyd, & Roberson, 2015) and do not always satisfy the "disability" criteria under disability discrimination laws. Therefore, even postsecondary students whose LD accommodations started in childhood may not be covered anymore when they reach the stage of taking the COMLEX.

Overdiagnosis of LD in postsecondary populations?

Research indicates that many college students classified as having a LD do not actually meet professionally recognized criteria for diagnosis of this disorder (e.g., Sparks & Lovett, 2009a, 2009b; Weis, Erickson & Till, 2017). In these studies, a large proportion of diagnosed students failed to meet various agreed upon criteria for diagnosis of LD, and the majority had no normative academic weakness relative to their same-aged peers. This mirrors the findings of a recent review of research by Harrison (2017), who concluded that a majority of clinicians undertaking evaluations for diagnosis of LD in postsecondary environments fail to follow any accepted diagnostic guidelines, with most using flexible or nonexistent criteria when making a LD diagnosis.

In adults, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; APA, 2013) provides authoritative clinical criteria, and the Americans with Disabilities Act (as amended; ADAAA, 2008) provides similar clear guidance for determining whether a condition (such as LD) causes sufficient functional impairment to qualify as a disability in the legal sense and lead to an entitlement to accommodations. Both the DSM-5 and ADAAA require below-average skills relative to a general population (or age-peer) standard. Hence, it is reasonable to characterize the large number of LD-diagnosed college students who do not meet these standards as a phenomenon of overdiagnosis.

This overdiagnosis—and Harrison's (2017) finding of flexible standards on the part of diagnosticians—may be due, in part, to clinicians adopting an advocacy role when completing disability assessments (Harrison, Lovett, & Gordon, 2013). Additionally, despite studies showing how easily postsecondary students can believably feign LD (e.g., Harrison & Edwards, 2008), 71% of clinicians surveyed by Harrison et al. (2013) were unaware of the research showing that clinicians are poor at detecting symptom

exaggeration when it occurs. Harrison et al. estimated that less than 10% of clinicians conducting postsecondary disability evaluations include specialized measures of test-taking engagement (performance validity tests, or PVTs; see, e.g., Larrabee, 2014a) in their assessments.

Another factor that may contribute to the overdiagnosis of LD in postsecondary students has to do with the norms used on achievement tests. Under U.S. federal laws, disabilities involve substantial limitations relative to an average person, general population standard. In addition, the DSM-5 (APA, 2013) clearly states that LD involves academic skills that are "substantially and quantifiably below those expected for the individual's *chronological age*" (p. 67, emphasis added). However, on many achievement tests (e.g., the Woodcock-Johnson Tests of Achievement (WJ-Ach; Schrank, Mather, & McGrew, 2014), grade-based norms are available as well. On some tests (e.g., the Nelson-Denny Reading Test; Brown, Fishco, & Hanna, 1993), *only* grade-based norms are provided. Since only a higher ability, more academically skilled subset of the population progresses to the highest levels of education, grade-based norms will become increasingly unrepresentative of the general population as the education level of the client increases. Consider an achievement test that was standardized in 2008; less than 30% of the U.S. adult population had a college degree at that point in time, and over 10% did not even have a high school diploma (United States Census Bureau, 2010). Comparing a student to norms of graduating college seniors (as are available on the Nelson-Denny Reading Test [NDRT]; Brown, et al., 1993), or to grade 17.9 norms (as are available on the WJ-Ach [Schrank, Mather, et al., 2014]) would lead to a severe underestimate of the student's academic skills relative to the average person in the general population (see Brooks, Sherman, Iverson, Slick, & Strauss, 2011, for further discussion of the role of the normative sample in interpreting scores).

Present study

Despite a number of studies having investigated whether college students with LD diagnoses actually meet formal criteria for deficits in academic skills relative to the general population (or same-age) peers, we were unable to find any studies specifically examining LD documentation in the context of accommodations requests on high-stakes postsecondary standardized tests. The present study reviewed 103 consecutive applications sent to the NBOME requesting disability accommodations on one or more parts of the COMLEX exam. These applicants reported that their need for accommodation was based on a diagnosis of LD, with or without comorbid conditions. Their documentation included a completed application form as well as reports from diagnostic evaluations and supplemental information (e.g., evidence of prior disability accommodations, academic transcripts).

Our research questions focused on the nature and accuracy of the LD diagnosis, as judged against official clinical standards (from the DSM) for diagnosis and legal standards (from the ADAAA) for disability determination. More specifically, we wished to assess objectively the documentation provided to answer the following questions:

1. How many applicants provided evidence that they met official DSM criteria for diagnosis of a LD?
2. How many requested extended testing time, and if so, how much? Were any differences found in scores between those requesting 50% as compared with 100% extra time?
3. What was the basis that evaluators used to make a LD diagnosis? How many LD diagnoses were specifically given based on a discrepancy between an IQ score and an achievement score (rather than based on absolute deficits in achievement)?
4. How often did clinicians employ grade-based norms as a means of calculating achievement test scores and thereby determining academic impairment? And finally,
5. How often were PVTs employed in such evaluations?

Method

Procedure

NBOME staff identified 103 consecutive accommodation requests submitted on the basis of LD in the years 2013 through 2016. Each accommodation request consisted of a completed application form, along with supportive documentation that typically included diagnostic evaluation reports, any evidence of eligibility for accommodations in prior settings, and letters from various professionals and others in support of the request. Some applicants also submitted academic transcripts, K–12 school records, and/or score reports from past standardized tests. Two of the three reviewers were licensed/registered psychologists and the third had graduate training and over 20 years of experience as a disability consultant.

Each file review was conducted by reading the application and supportive documentation, and completing a standardized review form. Three types of information were recorded on the form: (a) general information about the applicant and the requested accommodation(s), (b) the rationale used by the diagnostician to support an LD diagnosis, and (c) diagnostic test score data. No information was provided by the applicants with respect to ethnicity or socio-economic status. Standard scores ($M = 100$, $SD = 15$) and percentiles were recorded for performance on subtests and composites of the third or fourth editions of the Wechsler Adult Intelligence Scale (WAIS-III/WAIS-IV; Wechsler, 1997, 2008), the third or fourth editions of the Woodcock-Johnson battery (WJ-III/WJ-IV; Schrank, Mather, et al., 2014; Schrank, McGrew, & Mather, 2014; Woodcock, McGrew, & Mather, 2001), and the second or third editions of the Wechsler Individual Achievement Test (WIAT-II/WIAT-III; Wechsler, 2001, 2009). In addition, any available scores (number correct, separate "raw" scores, scale scores [$M = 200$, $SD = 25$], clinician-transformed standard scores, percentiles, and/or grade-equivalent scores) from the NDRT (Brown et al., 1993) were recorded. Finally, since research has found that a substantial minority of postsecondary students being evaluated for LD exaggerate their symptoms (e.g., Sullivan, May, & Galbally, 2007), and that PVTs are helpful in detecting such exaggeration, reviewers also recorded whether diagnosticians used any of these tools.

Interrater reliability

The three documentation reviewers practiced coding ten sample cases to detect any problems with use of a draft review form. Discussion of these cases and subsequent slight revision of the form led to a final version which was used for the 103 files. A complete form is available from the first author upon request. Interrater reliability was determined using Fleiss's kappa coefficient on key variables. Agreement among raters was almost perfect (κs ranged from .80 to .99 for categorical variables). All discrepancies were resolved by reviewing the original documentation provided by the candidate.

Results

Of the 103 participants, 55 (54%) were men. Most applicants were in their late twenties ($M = 27.4$ years; $SD = 3.3$) and 13.9% had undergone their most recent assessment within the year in which they applied for accommodations. Only 39 of the applicants (38%) provided evidence that they had experienced significant symptoms of LD prior to age 18. This included not only individuals who were formally diagnosed with LD in childhood, but also those who, despite no childhood diagnosis, provided other historical documentation that supported symptom onset before age 18.

Prior academic accommodations were provided to applicants for many reasons, including LD. About one quarter (26.5%) of the applicants had received academic accommodations of some type in elementary school, 35.6% had first received academic accommodations some time during high school, and13.3% began first receiving accommodations at some time between ages 19 and 22. At older ages, 16.3% first received accommodations between ages 22 and 30, and eight (8.2%) had only started receiving accommodations on or after age 30.

Diagnoses given in assessment reports

All applicants indicated on their application form that the reason for needed accommodations was a LD (alone or in combination with other conditions); however, only 81 had actually been explicitly given a diagnosis of a particular learning disability (e.g., in reading). Thirteen had been given a diagnosis of ADHD. Nine had been given either nonspecific diagnoses (e.g., LD not otherwise specified, dyslexic tendencies, poor reading, cognitive disorder not otherwise specified, executive functioning disorder), or other diagnoses (e.g., developmental coordination disorder, generalized anxiety disorder).

Objective test scores

Table 1 provides the cognitive and achievement test scores reported in each of the 103 applicants' most recent assessment reports. Most commonly, students were given the Wechsler Adult Intelligence Scale (WAIS, either third or fourth edition) to evaluate cognitive abilities, with only 10% being given the Woodcock-Johnson Tests of Cognitive Abilities (WJ-IV COG; Schrank, McGrew, et al., 2014) as the sole measure of intellectual functioning, although selected subtests of the WJ-IV COG were also given as part of some other applicants' test batteries.

On average, applicants' Full Scale IQ scores on the WAIS fell within the high average range, exceeding approximately 79% of individuals in the general

Table 1. Diagnostic test scores reported in applicants' documentation.

	N	Mean	*SD*	Range	%≤85	%>85
WAIS						
Verbal Comp	94	120.7	12.6	93–150	0	100.0
Perceptual Reason	94	111.6	12.3	79–138	2.1	97.9
Working Memory	95	104.2	12.9	73–142	7.5	92.5
Processing Speed	94	98.1	13.7	73–127	21.7	78.3
FSIQ	92	112.9	10.8	80–138	1.1	98.9
GAI	31	122.8	11.5	99–148	0	100.0
WJ						
Letter Number Match	4	91.3	32.7	53–132	66.7	33.3
Concept Formation	11	113.0	12.1	95–125	0	100.0
Number Pattern Match	4	86.3	23.4	56–110	50.0	50.0
V A Learning	12	96.3	17.7	58–118	25.0	75.0
Story Recall	6	93.0	11.4	77–107	16.7	83.3
Anal Synthesis	8	113.9	10.5	91–126	0	100.0
Fluid Reason	10	114.2	16.3	88–139	0	100.0
General Intell. Ability	7	104.7	9.4	97–118	0	100.0
Processing Speed	16	88.3	9.1	78–107	50.0	50.0
WIAT						
Listening Comp	16	103.0	16.5	56–122	6.3	93.8
Oral Language	12	111.9	16.2	72–128	8.3	91.7
Sentence Comp	10	111.6	16.5	81–129	20.0	80.0
Essay Comp	15	111.7	13.8	91–142	0	100.0
Spelling	25	100.3	13.3	71–122	16.0	84.0
Written Expression	19	105.0	176	64–129	10.5	89.5
Reading Comp	27	104.7	19.2	55–150	14.8	85.2
Oral Reading Fluency	15	100.3	4.8	94–112	0	100.0
RC and Fluency	10	104.3	8.3	90–117	0	100.0
Word Reading	26	104.7	10.3	75–120	3.8	96.2
Pseudoword Decode	25	100.3	15	66–129	16.0	84.0
Basic Reading	13	101.8	12.6	82–129	15.4	84.6
Math P	20	112.5	13.5	75–136	5.0	95.0
Numerical Operations	25	111.8	11.2	92–131	0	100.0
Math	13	115.7	11.5	97–134	0	100.0
Math Addition	12	92.8	15.4	75–116	33.3	66.7
Math Subtraction	12	95.8	12.0	81–117	25.0	75.0
Math Multiplication	12	97.7	11.6	83–123	16.7	83.3
Math Fluency	10	94.9	11.0	84–117	20.0	80.0
WJ						
Reading Fluency	72	90.9	12.7	45–123	34.3	65.7
Letter Word ID	71	101.4	8.0	75–120	2.9	97.1
Sound Blend	11	106.4	13.4	88–125	0	100.0
Broad Reading	58	96.5	14.6	9–116	10.7	89.3
Reading Vocabulary	10	103.6	10.9	86–122	0	100
Word Attack	50	95.2	11.8	69–118	18.4	81.6
Basic Reading	31	96.3	7.6	77–113	6.9	93.1
Passage Comp	67	104.3	9.4	72–133	3.1	96.9
Reading Recall	3	95.7	7.2	91–104	0	100.0
Reading Comp	12	108.2	11.4	87–131	0	100.0
Spell Sound	13	93.5	12.9	75–121	23.1	76.9
Oral Comp	6	98.8	7.9	86–107	0	100.0
Oral Language	5	106.8	12.5	92–121	0	100.0
Sound Awareness	4	952.0	5.6	87–99	0	100.0
Understand Direct	8	104.3	20.0	76–126	25.0	75.0
Listen Comp	7	104.5	13.3	82–117	16.7	83.3
Calculation	66	110.7	9.5	87–132	0	100.0
Math Facts Fluency	68	94.7	14.0	59–124	21.2	78.8
Math Calc Skills	48	108.0	12.4	75–137	6.5	93.5
Applied Problems	48	108.1	10.4	76–127	6.4	93.6
Broad Math	42	106.8	11.6	74–134	4.9	95.1
Number Matrix	5	106.6	11.3	95–120	0	100.0
Math Problem Solv	2	113.5	17.7	101–126	0	100.0
Spelling	66	100.4	12.7	65–126	17.2	82.8
Sent Write Fluency	61	101.5	13.2	68–133	12.1	87.9
Writing Samples	50	110.3	12.7	77–148	4.1	95.9
Broad Written Lang	43	105.5	11.7	70-–124	7.1	92.9
Editing	4	80.3	46.1	12–113	25.0	75.0
Basic Write Skills	8	101.5	13.2	73–117	14.3	85.7
Written Expression	38	108.2	11.2	78–125	2.8	97.2
Academic Fluency	42	93.4	12.3	61–120	20.0	80.0
Nelson Denny	*N*	Mean	*SD*	Range	≤175	>175
VC Scaled Score	29	225.7	18.6	183–258	0	100.0
RC Scaled Score	63	202.5	18.2	160–249	8.1	91.9
RR Scaled Score	58	178.0	14.0	136–213	40.4	59.6

Note. WAIS = Wechsler Adult Intelligence Scale (3rd and 4th Editions); WJ = Woodcock-Johnson Test of Cognitive Abilities, (3rd or 4th Editions), and Woodcock-Johnson Test of Achievement, (3rd or 4th Editions); WIAT = Wechsler Individual Achievement Test; WRAT = Wide Range Achievement Test; CTONI 2 = Comprehensive Test of Nonverbal Intelligence 2. GE = grade equivalent norms, AE = age equivalent norms; VC = Vocabulary; RC = Reading Comprehension; RR = Reading Rate. Tests given to only 1 applicant are not reported.

[a]∼20% report norms used; only 1 used GE, 20 used AE. 13.% WIAT II; 12.9% WIAT III.

[b]8 clinicians used GE, 53 used AE, 16 did not report norms used.

population, whereas their measured General Intellectual Ability (GIA) on the WJ-IV COG fell in the average range, better than about 60% of their peers (Table 1). At the individual level, two applicants had FSIQ scores below 90 (one had a FSIQ of 80). Approximately one third of the participants (37) had IQ/GIA scores in the average range (90 to 109), an equal number fell in the high average range, and 23 (22%) had general intelligence scores in the superior range (120 to 129) or higher.

In general, applicants had higher scores on measures of verbal reasoning (superior range), high average perceptual reasoning, and average scores on the other two composite scores of the Wechsler scales (Working Memory Index and Processing Speed Index). Of note, 7.5% of applicants were said to have impairments (i.e., scores below the 16th percentile) in working memory, and 21.7% fell below this level on measures of processing speed. Those who were given the WJ-IV COG showed no deficits in overall cognitive skills (GIA), although half obtained scores below the 16th percentile on processing speed composite of this test.

Applicants' mean academic achievement scores also fell within the average to above average range and were relatively consistent across academic domains. Although clinicians primarily assessed academic achievement using either a version of the Wechsler Individual Achievement Test (WIAT-II or III) or a version of the WJ Tests of Achievement (WJ-III or WJ-IV), over half also used one or more of the NDRT subtests. Three reports explicitly diagnosed a LD based solely on intra-cognitive discrepancies (e.g., WAIS WMI significantly lower than VCI). Even though the mean academic fluency scores all fell within the average range, Table 1 shows that reading and math fluency subtests were the ones on which a larger percentage of applicants (20% on WIAT math fluency to 34.3% on WJ reading fluency) scored below the 16th percentile. When interpreting the meaning of obtained scores on the WJ, eight assessors reported using grade-based norms, 53 reported using age-based norms, and the remainder did not report which norms were employed.

By contrast, all clinicians used grade-based norms when documenting degree of impairment on the NDRT, typically employing the normative data for those at the end of a 4-year degree (71.0% of assessors used norms for grade 16, and 16.2% used norms for grades 14/15; the remainder used norms from lower grade levels, consistent with the applicant's grade level at the time of assessment). In addition, scale scores ($M = 200$, $SD = 25$) are available for the NDRT, "based on the pooled standardization sample from Grades 10, 11, and 12 and both two-year college classes and both lower division classes in the four-year institutions" (Brown et al., 1993, p. 14). Given that a large proportion of the population at the time of this tests' norming had not completed a four-year university program, and that upper-year postsecondary norms for this test consistently classify otherwise average individuals as being impaired, Harrison and Harrison (2019) recommended that these "pooled" scale scores be used when determining normative impairment in young adults. When the reported (grade-based) NDRT scores for these applicants were converted into these scale scores, the mean NDRT scores fell within the average range. For example, using grade-based norms on the NDRT the following percentage of applicants scored below or at the 16th percentile: for vocabulary, 13.3%, for comprehension, 50%, and for the one-minute reading rate, 74.6%. In contrast, far fewer scored at or below the 16th percentile when compared with the "pooled" scale score norms from the broad normative sample ($M = 200$, $SD = 25$): for vocabulary 0%, for comprehension, 7.9%, and for the one-minute reading rate, 41.4%. Similar decreases were seen in the percentage scoring at or below the 25th percentile when employing grade versus general population norms.

Evidence supporting diagnoses

Table 2 presents the percentage of students with evidence supporting their learning disability diagnosis. In 64 of the assessment reports, a discrepancy between intelligence and achievement was employed as the justification for the diagnosis given; 40 of these reports identified discrepancies between IQ and the grade-based NDRT scores as proof of impairment (despite IQ being based on age norms and NDRT scores being based on very different grade norms, virtually guaranteeing a discrepancy). It was therefore of interest to investigate whether the actual data even met criteria for a discrepancy-based diagnostic model (i.e., DSM-IV), and how many students demonstrated normative impairments in academic performance.

Even though only 81 of the 103 applicants had been given a formal diagnosis of some type of LD, all 103 had requested an extended time accommodation, so all 103 applicants would be expected to have demonstrated some type of impairment in functioning on an academic test (including diagnostic achievement tests). Table 2 shows the number of students for whom assessment data demonstrated that any type of

Table 2. Evidence supporting diagnosis based on published criteria.

	Yes	No			
Discrepancy used in diagnosis	64	39			
Test Data			Nelson Denny		
FSIQ-achievement discrepancy	Broad Skills (%)	Fluency (%)	Vocab	Comp	Reading Rate
1.0 *SD*	18.4	15.5	1.0	5.8	6.8
1.5 *SD*	10.7	15.5	1.9	7.8	10.7
2.0 *SD*	6.8	20.4	0	3.9	21.4
DSM-IV criteria	5.8	16.5			
DSM-5 (standard)	1.0	2.9			
DSM-5 (lenient)	2.9	11.7			

Note. Broad skills refers to a composite score from either the Woodcock-Johnson Tests of Achievement (WJ) or the Wechsler Individual Achievement Test (WIAT). Fluency scores similarly came from the WJ or WIAT.

discrepancy-based diagnostic criteria were met, regardless of whether the achievement score was average or better. The majority of students' performance on various measures of academic achievement did not support their need for accommodations on the COMLEX. For example, only 18.4% showed a 1 standard deviation (*SD*) discrepancy between their cognitive ability score (e.g., FSIQ) and any composite achievement score, 10.7% of students whose clinical reports recommended additional time on exams showed a 1.5 *SD* ability-achievement discrepancy, and 6.8% showed a 2 *SD* or more discrepancy; taking childhood onset into account, only 5.8% met the DSM-IV criteria for diagnosis of LD. Only one student evidenced a normative deficit in any achievement domain necessary to satisfy the standard (i.e., sum of squares [SS] ≤ 78) DSM-5 criteria for diagnosis of LD. Even when a more lenient (i.e., $SS \leq 85$) cutoff was applied, only 2.9% of students met DSM-5 criteria for impairment based on composite academic scores.

A slightly higher percentage of students met objective criteria based on WJ/WIAT academic fluency subtest scores. For instance, 15.5% of students whose clinicians recommended additional time on exams showed either a 1.0 or 1.5 *SD* discrepancy between their cognitive ability score and their reading, math, or writing fluency subtest scores, and 20.4% had a 2 *SD* discrepancy. Similarly, 11.7% of students met the DSM-5 criteria for reading, math, or writing fluency using the more lenient (i.e., $SS \leq 85$) cutoff, while only 2.9% met the stricter requirement (i.e., $SS \leq 78$). A slightly larger group (16.5%) met the DSM-IV criteria when employing some measure of academic fluency.

Although the test manual warns users that obtained scores on the NDRT are not to be used for purposes of diagnosis, and that the one minute reading rate subtest has weak reliability (0.68; Brown et al., 1993), an LD diagnosis was made for 40/103 applicants based exclusively on a discrepancy between some measure of IQ and one or more grade-based scores from this test. When students' scale scores from the NDRT were interpreted relative to most other individuals in the general population ($M = 200$, $SD = 25$) rather than by grade level, almost no vocabulary scores were below levels predicted by intelligence and fewer than 10% of the timed reading comprehension scores were more than one standard deviation below measured IQ (Table 2). Interestingly, 21.4% of applicants performed 2 SD or more below expected on the one minute reading rate subtest (Table 2).

Evidence supporting accommodation requests

Just over half of the students had historical evidence supporting their current accommodation requests. For example, prior to requesting accommodations on the COMLEX, 43 students (42%) had never before been provided with accommodations on any *high stakes* examinations they had taken such as the SAT, ACT, or MCAT. By contrast, 60 (58%) reported having received some type of accommodations in an *educational* setting prior to college.

Of the 103 reports reviewed, only 8 used at least one performance validity measure, and one included a *symptom* validity test (which can detect noncredible reports of subjective symptoms). The Word Memory Test (Green, 2005) was the most commonly used measure. Generally, the applicants "passed" the validity measures; however, one clinician who used a PVT and obtained evidence of exaggeration (i.e., the client failed the measure) noted that he had given the test to the client again after confronting the client about the failure and the client passed on the second attempt, which was interpreted as evidence that the client was not feigning or exaggerating symptoms.

All candidates seeking accommodations requested extra time: five applicants requested extra time but did not specify the amount required; two requested unlimited time; two asked for 25%; one for 33%; the

majority (67) requested 50%; and a quarter (26) requested double time. We compared those seeking up to 50% extra time with those seeking 100% extra time on the following achievement/fluency tests: WAIS FSIQ, WIAT Basic Reading Scale Score, Woodcock Johnson Sentence Reading Fluency, and Woodcock Johnson Broad Reading Fluency scores using a *t*-test for independent samples. No statistically significant differences were found between the achievement/fluency scores reported for those requesting 50% and 100% extra time, all *p*s ranging from .19 to .91 on these comparisons.

Discussion

To make appropriate decisions regarding accommodations eligibility, testing entities often rely on the self-report of candidates and the opinions of diagnosticians. Our findings raise some concerns about accepting at face value either of these sources of evidence. Indeed, despite requesting accommodations on the basis of a LD, we found that almost one quarter of applicants in our sample had never even received a formal diagnosis of a learning disability, at least not in their most recent psycho-educational evaluation, and often nowhere in their file. There are some proposed accommodation application guidelines that suggest taking students' self-reports as primary documentation of a disability (see Lovett, Nelson, & Lindstrom, 2015, for discussion), and the present study's finding casts doubt on such proposals. It is likely that the term "learning disability" is particularly apt to be misunderstood by applicants, since the term is used colloquially for a variety of developmental and neurologic problems, and so careful inspection of an official evaluation report is the first step toward establishing the presence of a bona fide LD.

The majority of candidates applying for accommodation due to a LD did not provide data supporting impairment in any relevant academic abilities compared to most other people in the general population. Indeed, using a lenient interpretation of DSM-5 diagnostic criteria, we found that only 2.9% met the criteria for diagnosis of LD in an overall academic skill area, and fewer than 12% met these same criteria for academic fluency impairments relative to the average person. Even applying the more lenient DSM-IV criteria resulted in only 5.8% (broad skills) and 16.5% (fluency) meeting diagnostic standards. This finding is similar to that reported by Harrison, Nichols, and Larochette (2008), who found that only 5% of postsecondary students seeking academic accommodations for LD in Canadian colleges provided comprehensive documentation that supported their disability diagnosis. Examining historical information, one quarter of the total sample reported receiving academic accommodations before age 12, and just over half received some type of school-based accommodations prior to age 19. These accommodations were not always given for LD (e.g., sometimes for other difficulties such as test anxiety or simply because it improved their academic performance). Notably, almost 40% of applicants were first given accommodations after age 18 and one quarter after age 22. Given that LD onsets in childhood and significantly affects academic functioning, the finding that such a large proportion had been able to function without accommodations until undergraduate or graduate school raises further doubts about the validity of the diagnoses.

While the majority of applicants requested 50% extra time, many requested double time. It was therefore interesting that no statistically significant differences were found between these two groups in terms of academic or fluency scores. This finding suggests that the extra time recommendations made by clinicians are likely not based on severity of cognitive or academic deficits as seen in actual diagnostic test performance. Instead, the amount of extended time recommended might be based on evaluator or client preference, or simple habit of what has been recommended in the past.

We found wide-ranging general ability (intelligence) levels; verbal comprehension scores were typically highest in a profile on the Wechsler scales and processing speed scores were typically lowest. Although the lower processing speed scores may seem consistent with requested extended time accommodations, the average processing speed score was still well within the average range for the general population (WAIS PSI = 98.1), and gifted individuals often have relatively lower processing speed scores (see e.g., the WAIS-IV manual, Wechsler, 2008). Meanwhile, the wide-ranging general ability levels of these applicants suggest that medical students who do not have above average intelligence may be struggling on physician licensure exams for that very reason; above average general ability may be needed for success in medical school and beyond, and just over half of the applicants had FSIQ scores in this range.

We found students' mean scores on multiskill achievement tests (the Woodcock-Johnson and WIAT) were always in the average range or above, with an exception of one WJ-III/WJ-IV subtest (Editing) that was only given to four participants. The

vast majority of the participants obtained standard scores of at least 86 on the timed diagnostic achievement subtests (i.e., fluency tests). Given these findings, it is not surprising that fewer than one fifth of participants met even the DSM-IV criteria for LD (fluency), with even fewer meeting the DSM-5 criteria, especially under the recommended strict version requiring academic skills at least 1.5 *SD* below the average. More students—but still only a minority—had an IQ-Achievement discrepancy of some type regardless of whether normative impairment was present. These findings are consistent with past research (Lovett & Sparks, 2010; Sparks & Lovett, 2009a,b; Weis et al., 2012) that found that both college and high-IQ samples with LD diagnoses generally have academic skills in the average range or above, and generally fail to meet formal LD criteria. The present study adds, then, to the evidence supporting Harrison's (2017) conclusions that diagnosticians apply flexible criteria for LD diagnosis rather than adhering to official guidelines.

As with WJ and WIAT achievement scores, most students' NDRT scores were in the average range when general population comparisons were applied (using the scale scores of $M = 200$, $SD = 25$). Fewer than 10% of the NDRT timed reading comprehension scale scores were 175 or below. Clinicians generally used the highest grade level norms (Grade 16 Spring) for scoring the NDRT, thus invalidating the scores for purposes of LD diagnosis. The current DSM guidelines clearly state that *age* expectations are the standard against which to judge academic skill deficits, and the legal ADAAA standard for disability is the general population, most of whom are not college graduates. To examine the difference between Grade 16 Spring norms and the general population proxy (scale scores) as recommended by Harrison and Harrison (2019), we recorded which applicants were reported by their evaluators to have NDRT scores that were at least 1 *SD* below the mean (generally based on Grade 16 Spring norms), and which applicants had such low scores compared to most other people (i.e., the general population proxy, $M = 200$, $SD = 25$). The differences were quite large; for instance, whereas half of participants were reported to have low ($\leq$16th percentile) NDRT timed comprehension scores when Grade 16 Spring norms were employed, fewer than 8% of participants met that standard when using the pooled norms. These differences were consistent with those found by Harrison and Harrison (2019), with prior studies comparing age and grade-based norms on the WJ battery (Cressman & Liljequist, 2014; Giovingo, Proctor, & Prevatt, 2005), and with known psychometric errors caused by applying normative data derived from high-functioning samples (e.g., Brooks et al., 2011).

Similarly, while the timed reading rate measure of the NDRT may seem a good proxy for evaluating reading speed, it suffers from two significant problems. First, it has weak reliability ($r = 0.68$) and questionable predictive validity. Indeed, after finding that this measure correlated only weakly with other age-normed measures of reading speed as well as with the NDRT Reading Comprehension score for any given student, Lewandowski, Codding, Kleinmann, and Tucker (2003) concluded that the NDRT reading rate measure could not be trusted to make decisions about relative reading proficiency.

Second, the Reading Rate subtest requires only that the examinee indicate how much they have read in one minute. Given that slowing down reading speed is one of the main strategies employed by those feigning a reading disorder (Harrison et al., 2008), and there is no way to verify how quickly the client can actually read, it would be relatively easy for students seeking extra time accommodations to slow down their reading or indicate a low level of completion after one minute. Consistent with the conclusions of both Smith (1998) and Ward-Murray (1998), one should therefore not employ the Reading Rate scores from the NDRT as the sole indicator of impaired reading fluency. Apart from being unreliable, it is also a subjective measure that is face valid and therefore easy to feign if one is motivated to demonstrate the need for extra time accommodations.

Relatedly, since LD evaluations are often done for the explicit and open purpose of documenting accommodations needs, thus giving students a possible incentive to exaggerate any impairment, we were interested in whether evaluators used PVTs to measure client effort/motivation. Fewer than 10% of evaluators examined low effort/motivation as a potential contributor to any low scores obtained on cognitive or achievement tests. Furthermore, even when a positive finding of symptom exaggeration was obtained, the diagnostician in question dismissed this evidence and readministered the same PVT.

The fact that so few clinicians even consider symptom exaggeration (and may even ignore positive evidence when such behavior occurs) is troubling. Research has shown clearly that students motivated to feign LD can produce a profile of below-average scores on measures of academic skills, fluency, and processing speed that are equal to or lower those found in persons with genuine LD (Harrison,

Edwards, & Parker, 2008; Lindstrom, Coleman, Thomassin, Southall, & Lindstrom, 2011). As such, it is likely that students motivated to receive extra time accommodations could produce test scores that would be interpreted as indicating a substantial impairment in academic functioning and qualify for test accommodations if performance validity measures were not utilized. Passing a *symptom* validity test (SVT) such as those embedded in certain self-report personality questionnaires during a LD assessment is not proof of effortful performance because those SVTs may only be sampling a tendency to exaggerate psychiatric disorders such as schizophrenia (Sullivan et al., 2007).

Implications for practice

Our findings suggest that when issuing LD diagnoses, especially formal DSM-5 diagnoses, evaluators should follow the official guidelines for that diagnostic category, and explain the evidence that shows the client met those guidelines. Although use of the NDRT may be defensible, given the paucity of alternative timed measures of reading comprehension using multiple-choice questions, evaluators should not base diagnostic judgments on scores from Grade 14–16 norms, which radically underestimate clients' skill levels relative to the general population, or on the reading rate subtest which is both unreliable and easily manipulated (see also Harrison & Harrison, 2019, for an extended discussion of these concerns). Instead, for a general population proxy group, Harrison and Harrison (2019) recommend that clinicians employ either Grade 12 or the "pooled" Scale Score norms of the NDRT for disability determination purposes. Finally, evaluators should strongly consider adding formal measures of performance validity to their standard LD diagnostic battery, particularly when there is an explicit question at the time of referral concerning accommodations eligibility.

Our findings are even more important for testing entities (e.g., medical schools, state licensure boards, independent testing agencies) considering accommodation requests from medical students and physicians. Unfortunately, our findings suggest that testing entities cannot trust applicants' self-reports of having LD to mean that a formal diagnosis was ever made, and that even when a diagnosis has been made, the applicant will often not meet the official diagnostic criteria for the disorder. Moreover, the present results reinforce the statement found in the DSM-5 that a clinical diagnosis of a disorder does not imply that someone meets the legal test for a disability (American Psychiatric Association, 2013, p. 25). In general, testing entities may find that many applicants reporting LD will have academic skills in the average range, thus failing to meet the "substantial limitation" threshold in disability law. Finally, even when low achievement scores are found in an applicant's file, they may be due to the clinician's use of grade-based norms or the clinician's failure to consider motivation to obtain accommodations as a potential cause of any low scores. In short, if testing entities wish to ensure that accommodations are provided to those applicants who truly meet official clinical and legal standards then they cannot take applicants' self-reports or diagnosticians' judgments at face value without further inquiry.

Limitations and directions for future research

The main limitation is the generalizability of results. Since we examined files from *all* students applying for LD accommodations on the COMLEX between 2013 and 2016, our results likely apply well to other osteopathic medical students with LD diagnoses. However, it is unclear how broadly our results apply to other groups of students applying for testing accommodations. As we noted above, college students with LD diagnoses seem similar to our sample in some ways (see Sparks & Lovett, 2009a), but groups of students requesting LD accommodations in other allied health professions, for instance, may yield different profiles. Indeed, such a result is suggested by Weis, Sykes, and Unadkat (2012), who found that even college type (e.g., community college, public university, private liberal arts college) substantially affected what proportion of students would meet different LD diagnostic criteria sets. We would therefore encourage future investigators to extend our methodology to related populations (e.g., allopathic medical schools, nursing programs).

A second potential limitation pertains specifically to our analyses concerning what proportion of our sample met criteria for LD under the DSM-5 criteria. As described above, we applied two different interpretations of the DSM-5 criteria, but both interpretations required that students have academic skills that were below-average, relative to age expectations. The DSM-5 text accompanying the criteria does contain a passage suggesting that in intellectually gifted students, "apparently adequate academic functioning" can be sustained even in the presence of LD, at least in certain settings (American Psychiatric Association, 2013, p. 69). It is possible that diagnosticians are using this

passage when making a DSM-5 diagnosis of LD in medical students with high IQs. However, even when diagnostic criteria were applied that rely only on relative impairment (e.g., DSM-IV, which allows for a disability diagnosis when a discrepancy between IQ and achievement is present regardless of normative impairment), few of these diagnosed students met these discrepancy criteria.

Conclusions

The present study found that many applicants seeking accommodations on licensure tests do not appear to meet relevant clinical and legal standards of disability, at least on the basis of the documentation that they have submitted. To be clear, there is no doubt that some medical students meet criteria for learning disabilities, and should be provided with appropriate accommodations matched to their functional limitations in test-taking settings. However, medical students applying for accommodations due to LD often fail to show sufficient evidence of a" substantial limitation" relative to most other individuals in the general population (as required under the ADAAA) and often fail to meet accepted diagnostic criteria. Indeed, contrary to what is required to make this diagnosis, the majority of applicants in this study demonstrated satisfactory academic skills. Further, the amount of extra test-taking time recommended in a report does not appear to be associated with an applicant's degree of measured functional impairment. We recommend that evaluators adhere to formal criteria for diagnosis, and that testing entities carefully review documentation to ensure that evaluators have done so. Such procedures will ensure that accommodations are targeted toward students who truly need them, while maintaining test integrity and fairness for all students.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Americans with Disabilities Act Amendments Act (ADAAA), 42 U.S.C. §12101 et seq, Pub. L. No. 110-325. (2008).

Brooks, B., Sherman, E., Iverson, G., Slick, D. J., & Strauss, E. (2011). Psychometric foundations for the interpretation of neuropsychological test results. In M. R. Schoenberg & J. G. Scott (Eds.), *The little black book of neuropsychology: A syndrome-based approach* (pp. 893–922). New York, NY: Springer.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson Denny Reading Test: Manual for scoring and interpretation, Forms G and H*. Chicago, IL: Riverside.

Cortiella, C. (2013). *Diplomas at risk: A critical look at the graduation rate of students with learning disabilities*. New York, NY: The National Center for Learning Disabilities.

Cortiella, C., & Horowitz, S. H. (2014). *The state of learning disabilities: Facts, trends and emerging issues*. New York, NY: National Center for Learning Disabilities.

Cressman, M. N., & Liljequist, L. (2014). The effect of grade norms in college students: Using the Woodcock–Johnson III Tests of Achievement. *Journal of Learning Disabilities*, *47*(3), 271–278. doi:10.1177/0022219412461052

Giovingo, L. K., Proctor, B. E., & Prevatt, F. (2005). Use of grade-based norms versus age-based norms in psychoeducational assessment for a college population. *Journal of Learning Disabilities*, *38*(1), 79–85. doi:10.1177/00222194050380010601

Green, P. (2005). *Manual for the Word Memory Test for Windows* (Rev. ed.). Edmonton, Alberta, Canada: Green's Publishing.

Harrison, A. G. (2017). Clinical, ethical, and forensic implications of a flexible threshold for LD and ADHD in postsecondary settings. *Psychological Injury and Law*, *10*(2), 138–150. doi:10.1007/s12207-017-9291-9

Harrison, A. G., & Edwards, M. J. (2008). How easily can post-secondary students feign symptoms of reading disabilities? *Archives of Clinical Neuropsychology*, *23*, 745–745.

Harrison, A. G., Edwards, M. J., & Parker, K. C. (2008). Identifying students feigning dyslexia: Preliminary findings and strategies for detection. *Dyslexia*, *14*, 228–246. doi:10.1002/dys.366

Harrison, A. G., & Harrison, K. A. (2019). A critical analysis of the Nelson Denny reading test as a method of identifying reading impairment in adults. *Psychological Injury and Law*. Online First Publication, doi:10.1007/s12207-019-09340-2

Harrison, A. G., Lovett, B., & Gordon, M. (2013). Documenting disabilities in postsecondary settings: Diagnosticians' understanding of legal regulations and diagnostic standards. *Canadian Journal of School Psychology*, *28*(4), 303–322. doi:10.1177/0829573513508527

Harrison, A. G., Nichols, E., & Larochette, A. (2008). Investigating the quality of learning disability documentation provided by students in higher education. *Canadian Journal of School Psychology*, *23*(2), 161–174. doi:10.1177/0829573507312051

Harrison, A. G., & Wolforth, J. (2012). Findings from a pan-Canadian survey of disability services providers in postsecondary education. *International Journal of Disability, Community & Rehabilitation*, *11*. Retrieved from: http://www.ijdcr.ca/VOL11_01/articles/harrison.shtml

Kortering, L. J., & Braziel, P. M. (1999). Staying in school: The perspective of ninth-grade students. *Remedial and Special Education*, *20*(2), 106–113. doi:10.1177/074193259902000208

Larrabee, G. J. (2014). Test validity and performance validity: Considerations in providing a framework for development of an ability-focused neuropsychological test battery. *Archives of Clinical Neuropsychology*, *29*(7), 695–714. doi:10.1093/arclin/acu049

Lewandowski, L. J., Codding, R. S., Kleinmann, A. E., & Tucker, K. L. (2003). Assessment of reading rate in

postsecondary students. *Journal of Psychoeducational Assessment*, *21*(2), 134–144. doi:10.1177/0734282912462693

Lindstrom, W., Coleman, C., Thomassin, K., Southall, C., & Lindstrom, J. (2011). Simulated dyslexia in postsecondary students: Description and detection using embedded validity indicators. *The Clinical Neuropsychologist*, *25*(2), 302–322. doi:10.1080/13854046.2010.537280

Lindstrom, W., & Lindstrom, J. (2017). College admissions tests and LD and ADHD documentation guidelines: Consistency with emerging legal guidance. *Journal of Disability Policy Studies*, *28*(1), 32–42. doi:10.1177/1044207317696261

Lovett, B. J., Gordon, M., & Lewandowski, L. J. (2016). Legal conceptions of impairment: Implications for the assessment of psychiatric disabilities. In S. Goldstein & J. A. Naglieri (Eds.), *Assessing impairment: From theory to practice* (2nd ed., pp. 125–139). New York, NY: Springer.

Lovett, B. J., & Lewandowski, L. J. (2015). *Testing accommodations for students with disabilities: Research-based practice*. Washington, DC: American Psychological Association Press.

Lovett, B. J., Nelson, J. M., & Lindstrom, W. (2015). Documenting hidden disabilities in higher education: Analysis of recent guidance from the association on higher education and disability (AHEAD). *Journal of Disability Policy Studies*, *26*(1), 44–53. doi:10.1177/1044207314533383

Lovett, B. J., & Sparks, R. L. (2010). Exploring the diagnosis of "Gifted/LD": Characterizing postsecondary students with learning disability diagnoses at different IQ levels. *Journal of Psychoeducational Assessment*, *28*(2), 91–101. doi:10.1177/0734282909341019

Maki, K. E., Floyd, R. G., & Roberson, T. (2015). State learning disability eligibility criteria: A comprehensive review. *School Psychology Quarterly*, *30*(4), 457. doi:10.1037/spq0000109

National Board of Osteopathic Medical Examiners. (2018). *COMLEX-USA Bulletin of information, 2018-2019*. Available at www.nbome.org.

Roberts, B. L. (2012). Beyond psychometric evaluation of the student—task determinants of accommodation: Why students with learning disabilities may not need to be accommodated. *Canadian Journal of School Psychology*, *27*(1), 72–80. doi:10.1177/0829573512437171

Rojewski, J. W. (1999). Key components of model transition services for students with learning disabilities. *Learning Disability Quarterly*, *15*(2), 135–150. doi:10.2307/1511015

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV Tests of Achievement*. Rolling Meadows, IL: Riverside.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV Tests of Cognitive Abilities*. Rolling Meadows, IL: Riverside.

Smith, D. (1998). Test review of the Nelson Denny Reading Test Forms G & H. In J. Impara & B. Plake (Eds.), *The Thirteenth Mental Measurements Yearbook (Buros Mental Measurements Yearbooks)* (pp. 685–686). Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska Press.

Sparks, R. L., & Lovett, B. J. (2009). College students with learning disability diagnoses: Who are they and how do they perform? *Journal of Learning Disabilities*, *42*(6), 494–510. doi:10.1177/0022219409338746

Sparks, R. L., & Lovett, B. J. (2009). Objective criteria for classification of postsecondary students as learning disabled: Effects on prevalence rates and group characteristics. *Journal of Learning Disabilities*, *42*(3), 230–239. doi:10.1177/0022219408331040

Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, *14*(3), 189–207. doi:10.1080/09084280701509083

U.S. Government Accountability Office. (2011). *Higher education and disability: Improved federal enforcement needed to better protect students' rights to testing accommodations (GAO-12–40)*. Washington, DC: Author.

United States Census Bureau. (2010). *Statistical abstract of the United States: 2010*. Retrieved from: https://www.census.gov/library/publications/2009/compendia/statab/129ed.html

Ward-Murray, M. (1998). Test review of the Nelson Denny Reading Test Forms G & H. In J. Impara & B. Plake (Eds.), *The Thirteenth Mental Measurements Yearbook (Buros Mental Measurements Yearbooks)* (pp. 683–685). Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska Press.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). San Antonio, TX: Pearson.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Pearson.

Wechsler, D. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio, TX: Pearson.

Weis, R., Erickson, C. P., & Till, C. H. (2017). When average is not good enough: Students with learning disabilities at selective, private colleges. *Journal of Learning Disabilities*, *50*(6), 684–700. doi:10.1177/0022219416646706

Weis, R., Sykes, L., & Unadkat, D. (2012). Qualitative differences in learning disabilities across postsecondary institutions. *Journal of Learning Disabilities*, *45*(6), 491–502. doi:10.1177/0022219411400747

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Rolling Meadows, IL: Riverside.

Yell, M. L. (2019). *The law and special education* (5th ed.). Boston, MA: Pearson.