Journal of Learning Disabilities

"Passageless" Administration of the Nelson–Denny Reading Comprehension Test: Associations With IQ and Reading Skills

Rebecca E. Ready, Maheen F. Chaudhry, Kelly C. Schatz and Sarah Strazzullo

"Passageless" Administration of the Nelson–Denny Reading Comprehension Test: Associations With IQ and Reading Skills

Rebecca E. Ready, PhD,[1] Maheen F. Chaudhry, BS,[1] Kelly C. Schatz, BS,[1] and Sarah Strazzullo, BA[1]

Abstract

There are few tests that assess reading comprehension in adults, but these tests are needed for a comprehensive assessment of reading disorders (RD). The *Nelson–Denny Reading Test* (NDRT) has a long-passage reading comprehension component that can be used with adolescents and adults. A problem with the NDRT is that reading comprehension test items can be answered correctly without reading the associated passage. The current study determined how IQ, verbal comprehension, and reading skills were associated with scores on a *passageless administration* of the NDRT. Results indicated that IQ, verbal comprehension, and broad reading skills were significantly associated with greater NDRT passageless scores. Results raise questions about the validity of the reading comprehension component of the NDRT and suggest that the test may have differential validity based on individual differences in vocabulary, general fund of knowledge, and broad reading skills.

Keywords

reading disability, learning disability, *Nelson–Denny Reading Test,* passageless, reading comprehension

Earlier identification of learning disabilities (LDs) as well as increased access to special education and enrichment services allow more students with LD to pursue higher education than ever before. Between 4.3% and 11.0% of students on 4-year college and university campuses are diagnosed with an LD, and rates may be as high as 23% at 2-year institutions and community colleges (American College Health Association, 2011; Educational Testing Service, Policy Information Center, 2007). LDs are most common in adult education programs, affecting 10% to 50% of participants (American College Health Association, 2011). Despite their increased participation in higher education and adult education programs, a majority of these students report that their LD has had a negative impact on their academic progress in the form of lower or incomplete grades or significant disruption of work.

Many adults with LDs seek academic accommodations to increase their chances of success in educational environments. Most adults must provide current documentation of their LD to be eligible for services. Many institutions and testing services require assessment data to be no more than 3 to 5 years old (Educational Testing Service, Office of Disability Policy, 2007; Law School Admission Council,

2012). Thus, there is a need for reliable and valid tests to measure LDs in persons older than 18.

Reading disorders (RDs) are the most common type of LD in childhood, and most children with an RD continue to struggle with reading as adults (Katz, Goldstein, & Beers, 2001; Shaywitz et al., 1999). Literacy problems are prevalent in adults with LDs (Patterson, 2008). Therefore, an essential component of a psychological evaluation for an adult LD is a test of reading comprehension. Currently, there is a vital lack of reliable and valid long-passage reading comprehension instruments for adults.

In most reading comprehension tasks, the test takers read a passage and then answer questions about what they have read. Response formats are often multiple choice. It is presumed that correct responses to reading comprehension test questions indicate reading comprehension ability, but this may not be true. A major threat to the validity of reading

[1]University of Massachusetts, Amherst, MA, USA

Corresponding Author:
Rebecca E. Ready, Department of Psychology, University of Massachusetts, 135 Hicks Way, Tobin Hall, Amherst, MA 01003, USA.
Email: ready@psych.umass.edu

comprehension tests is that the items can be answered correctly without having read the associated passage (Keenan & Betjemann, 2006). That is, in "passageless administration," test items are administered to test takers *without* the associated reading passages and correct responses are greater than would be expected by chance. On the *Gray Oral Reading Test–Third Edition* (GORT-3), 86% of items (56 out of 65) were answered at above-chance accuracy, and approximately 16 of these items had accuracy greater than 75%. The authors also demonstrated that passageless accuracy predicted performance on the GORT-3 items under normal administration conditions better than other indicators of reading proficiency. Children with a RD performed just as well on passage-independent items as typical children, whereas they performed worse than typical children on passage-dependent items (e.g., items that were not answered at rates greater than chance).

RD assessments with older adolescents and adults rely, in part, on the *Nelson–Denny Reading Test* (NDRT) because it is one of the only long-passage reading comprehension tests normed on adults (Brown, Fishco, & Hanna, 1993). Indeed, the NDRT is a critically important component of psychological evaluations for many adults. Scores on the NDRT are relied on by the Law School Admission Council to determine if an individual will be allowed extra time to take the LSAT (Law School Admission Council, 2012). The Educational Testing Service, which administers the SAT and GRE, lists the NDRT as an acceptable test of academic achievement, among several other tests, for determining LDs in adolescents and adults (Educational Testing Service, Office of Disability Policy, 2007).

The NDRT was normed for education rather than age, so it does not have an upper age limit for administration and can be used to assess adults from different age groups. The NDRT includes long passages for reading comprehension, which increase the face validity of the task, but there are several drawbacks to the NDRT that limit its utility. It was published in 1993, and updated normative data are not available (Brown et al., 1993). In addition, there are no reliability data and limited validity data in the manual. A major shortcoming is that the reading passages were extracted from high school and college texts, which means that many persons who take the NDRT will have familiarity with passage content prior to taking the test. Thus, some persons may have an unfair advantage in answering test items about the reading passages.

Indeed, recent evidence reveals that test takers can determine the correct answers to the NDRT at greater than chance accuracy (Coleman, Lindstrom, Nelson, Lindstrom, & Gregg, 2010). Coleman et al. (2010) administered passageless versions of the NDRT to 253 "typical" college students and 26 college students at risk for an LD or attention-deficit/hyperactivity disorder (ADHD). There are two versions of the NDRT, Form G and Form H. Half of the typical

students and all the at-risk students were administered Form G in passageless format, and half were administered Form H in the same fashion. Each reading comprehension item is multiple choice in format, with five response choices. Overall, typical college students responded correctly to between 44% and 47% of the items, significantly greater than chance accuracy of 20%. The at-risk student response rate was 41%. The NDRT reading passages cover humanities, social science, and science; all participant groups were more successful in responding to science questions (46.7% to 56.6%) than items from humanities passages (30.7% to 37.4%). NDRT reading comprehension items are evenly divided between *literal* and *interpretive*. Literal items assess explicit details of reading passages, whereas interpretive items require respondents to reason with information from the passages. Coleman et al. found that for Form H, literal and interpretive questions were answered with almost equal accuracy (~46%), whereas for Form G, interpretive items were answered correctly more often (47% to 54%) than literal items (30% to 33%). Overall, results of this study raised troubling concerns about the validity of the NDRT as a test of reading comprehension. The authors speculated that skill overestimation in NDRT scores may be particularly true for high-ability students who have a greater fund of general knowledge and/or superior verbal reasoning skills.

In the current study, individual differences in correct response rates to passageless administrations of the NDRT reading comprehension items were determined. Persons with greater estimated IQ and greater verbal comprehension skills were expected to be better at responding correctly in passageless administration of the NDRT. Persons with greater IQ may have a greater fund of general knowledge and/or superior skills in verbal reasoning that would aid in answering passageless items. Persons with superior reading skills also were expected to perform better on passageless comprehension items than persons with weaker skills. Superior reading skills may contribute to a better fund of knowledge and superior understanding of the test questions themselves. If predictions are correct, the NDRT may have differential validity based on these individual difference factors.

Method

Participants

Participants were 115 undergraduate students (31% male, 86% Caucasian; *M* age = 19.9 years, *SD* = 1.8); 61% were in their freshman or sophomore year of university. Participants were native English speakers and were recruited from psychology courses; participants received extra credit in a psychology course in exchange for participation. Participants who reported a past diagnosis of LD were oversampled to represent 25% of the sample.

Measures

Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV). The WAIS-IV comprises 10 core subtests, which compose four index scores: Verbal Comprehension Index (VCI), Perceptual Reasoning Index (PRI), Working Memory Index (WMI), and Processing Speed Index (PSI; Wechsler, Coalson, & Raiford, 2008). The Full-Scale Intelligence Quotient (FSIQ) is calculated from all 10 subtests. Internal consistencies range from .97 to .98 for FSIQ, from .87 to .98 for the four index scores, and from .73 to .96 for the subtests. The test–retest stability coefficients were highest for FSIQ (.94–.96) and VCI (.94–.95), followed by WMI (.82–.90), PRI (.80–.88), and PSI (.76–.89), and were lowest for the subtests (.51–.93). Interscorer agreement ranges from .91 to .99. Validity analysis showed that subtests within an index category were more highly correlated with each other than with subtests from a different index category. The FSIQ is highly correlated with academic achievement measures.

Woodcock–Johnson Tests of Achievement–Third Edition (WJ-Ach-III). Three subtests from the WJ-Ach-III that compose the Broad Reading cluster score were administered: Letter–Word Identification, Reading Fluency, and Passage Comprehension (Mather & Woodcock, 2001). Median reliabilities in the adult age range are .94 for Letter-Word Identification, .90 for Reading Fluency, and .88 for Passage Comprehension. The WJ-Ach-III demonstrates convergent validity with other tests of achievement (total scores correlate .65 to .79; Stetson, Stetson, & Sattler, 2001).

Passageless reading comprehension items from the NDRT. NDRT Forms G and H each comprise seven reading passages and 38 multiple-choice questions with five response options (Brown et al., 1993). Reading passages were derived from humanities, social sciences, and science sources. Half of the test items were developed to be literal and half to be interpretive. Reading Comprehension scores from Forms G and H, under normal administration conditions, correlate at .81. Validity data indicate that NDRT Reading Comprehension scores correlated .21 to .72 with final grades in reading and English courses; data were collected from three community colleges in California in 1993. Reliability data are not provided in the manual. In this study, participants were asked to complete the reading comprehension test items *without* reading the associated passages.

Procedure

All participants were assessed in individual appointments that lasted 2 to 2.5 hours. A trained research assistant described the study and obtained written informed consent. The research assistant administered the WAIS-IV core subtests according to standardized administration procedures. Participants were then administered the three WJ-III-Ach reading subtests. Finally, participants responded to passageless reading comprehension items from both the NDRT Forms G and H.

Results

Participants

The majority of participants reported that their physical (93%) and mental health (86%) was *good* or *very good*. Average household income ranged from $40,000 to $50,000. A quarter of the sample was previously diagnosed with an LD (*n* = 29); the age of LD diagnosis ranged from 4 to 20 years (*M* = 11.5, *SD* = 4.9). Of these persons with an LD diagnosis, 19 also had a diagnosis of ADHD; 11 other persons had a diagnosis of ADHD, without comorbid LD. Of the full sample, 29% reported a history of special education services at some point during their educational careers (*n* = 33). Of the 29 persons with an LD diagnosis, 24 reported special education services. Of the 11 persons with only ADHD, 2 reported special education services. Of the 19 persons with comorbid LD and ADHD diagnoses, 17 reported a history of special education services. In all, 21 persons reported current treatment for ADHD.

NDRT Passageless Scores

Respondents were able to correctly answer an average of 16 to 17 items on the NDRT Forms G (*M* = 16.44, *SD* = 3.67) and H (*M* = 16.97, *SD* = 3.99). Chance accuracy would be about 8 correct items; thus, similar to the findings of Coleman et al. (2010), correct response rates were at least double what would be expected by chance. Scores on the two NDRT forms were significantly correlated (Table 1).

Age was not significantly correlated with NDRT Form G and H scores (*r*s = −.10 and −.01, respectively, *p*s > .25); however, income was significantly correlated with the NDRT scores (*r*s = .30 and .19, respectively, *p*s < .05). Results of *t* tests indicated no gender differences between NDRT Form G and H passageless scores (*t*s = −0.76 and −0.56, respectively, *p*s > .45). There also were no significant differences between NDRT means for persons who were and were not diagnosed with an LD (*t*s = −0.39 and 0.63, respectively, *p*s > .35), who were or were not diagnosed with ADHD (*t*s = −0.47 and 0.26, respectively, *p*s > .60), who were or were not in special education at some point during their educational career (*t*s = −1.54 and −0.42, respectively, *p*s > .10), nor between persons who were and were not currently treated for ADHD (*t*s = 0.12 and 0.76, respectively, *p*s > .40).

Passageless NDRT Reading Comprehension, WAIS-IV, and WJ-III-Ach Scores

FSIQ and three of the four WAIS-IV index scores were significantly correlated with the NDRT scores. As predicted, the

Table 1. Correlations Between NDRT Forms G and H, WAIS-IV Index and IQ Scores, and WJ-III-Ach Broad Reading.

	1	2	3	4	5	6	7	8
1. NDRT Form G	—							
2. NDRT Form H	.54**	—						
3. WAIS-IV VCI	.49**	.42**	—					
4. WAIS-IV PRI	.14	.16	.48**	—				
5. WAIS-IV WMI	.28**	.29**	.35**	.45**	—			
6. WAIS-IV PSI	.28**	.23*	.18	.21*	.24*	—		
7. WAIS-IV FSIQ	.43**	.39**	.74**	.77**	.71**	.53**	—	
8. WJ-III-Ach B. Read.	.46**	.47**	.53**	.24**	.46**	.32**	.56**	—

Note: FSIQ = Full-Scale Intelligence Quotient; NDRT = *Nelson–Denny Reading Test*; PRI = Perceptual Reasoning Index; PSI = Processing Speed Index; VCI = Verbal Comprehension Index; WAIS-IV = *Wechsler Adult Intelligence Scale–Fourth Edition*; WJ-III-Ach B. Read. = *Woodcock–Johnson Tests of Achievement–Third Edition*, Broad Reading Composite; WMI = Working Memory Index.
*$p < .05$. **$p < .01$.

strongest WAIS-IV index association with the NDRT scores was with the VCI (Table 1); the FSIQ correlation with the NDRT scores is likely based largely on shared variance with the VCI since FSIQ and VCI are themselves highly correlated.

As predicted, the WJ-III-Ach Broad Reading score was significantly correlated with the NDRT score. In hierarchical regressions, associations between the NDRT scores and the VCI and Broad Reading were determined to ascertain if the VCI and Broad Reading composite accounted for unique variance in NDRT scores. VCI and Broad Reading were themselves correlated ($r = .53$, $p < .001$). Separate regressions were run with NDRT Forms G and H. Prior to entry into the regressions, VCI and Broad Reading were group-mean centered. Income was included in regressions because it was significantly associated with NDRT scores and because indicators of socioeconomic status have complex associations with indicators of IQ and achievement (e.g., Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). NDRT Form G was the first dependent variable; VCI and income were entered as the first independent variables in the regression, followed by Broad Reading. Results indicated that income and VCI accounted for significant variance in the NDRT Form G, and this was true even after Broad Reading was entered into the model. Broad Reading added significantly to the prediction of the NDRT Form G score (Table 2). Results followed the same pattern with NDRT Form H as the dependent variable; an exception was that income was no longer a significant independent variable in the regression model. Thus, skills assessed by the VCI and Broad Reading composites capture unique variance in passageless scores on the NDRT.

Literal Versus Interpretive NDRT Items

Half of the NDRT reading comprehension items are literal and half are interpretive. Similar to Coleman et al. (2010), we found that participants correctly answered significantly more interpretive ($M = 9.61$, $SD = 2.51$) than literal items

($M = 6.96$, $SD = 2.21$) for Form G ($t = 8.46$, $p < .001$). There were no significant differences in correct responses for interpretive ($M = 8.28$, $SD = 2.79$) and literal items ($M = 8.68$, $SD = 1.99$) for Form H.

Correlations between literal and interpretive items and WAIS-IV and WJ-III-Ach Broad Reading were calculated (Table 3). There were no systematic associations among IQ scores, Reading, and the NDRT items based on their literal versus interpretive scoring. That is, despite some correlations differing in magnitude between literal and interpretive items, none of the differences were significant using a t test after r to z transformation.

NDRT Items by Subject Area

On each form of the NDRT, 13 items are about humanities reading passages, 10 items pertain to social science, and 15 items assess comprehension of science passages. For humanities, 32% of From G ($M = 4.21$, $SD = 1.47$) and 29% of Form H ($M = 3.78$, $SD = 1.73$) items were answered correctly. For social science, 42% of Form G ($M = 4.18$, $SD = 1.54$) and 51% of Form H items ($M = 5.05$, $SD = 1.73$) were answered correctly. For science, 54% of Form G ($M = 8.04$, $SD = 2.07$) and 55% of Form H items ($M = 8.22$, $SD = 2.26$) were answered correctly. Thus, similar to Coleman et al. (2010), science questions were answered correctly more often than humanities items.

Correlations between NDRT items by content area and the WAIS-IV and WJ-III-Ach Broad Reading were calculated (Table 4). Again, VCI, FSIQ, and Broad Reading were most consistently correlated with NDRT scores and had the largest effect sizes.

VCI and Broad Reading Subtests and the NDRT Passageless Responses

Total NDRT scores, literal and interpretive scores, and content scores all were most consistently associated with the

Table 2. Hierarchical Regressions Predicting NDRT Scores From Income, WAIS-IV VCI, and WJ-III-Ach Broad Reading.

Dependent Variable	Independent Variables	B	SE	β	t
NDRT Form G	Block 1				
	Income	0.71	0.26	.23	2.75**
	VCI	0.14	0.03	.45	5.30**
	$R^2 = .30, F(2, 102) = 21.17**$				
	Block 2				
	Income	0.57	0.27	.19	2.15*
	VCI	0.11	0.03	.36	3.75**
	Broad Reading	0.10	0.05	.20	1.98*
	$\Delta R^2 = .03, \Delta F(1, 99) = 3.91*$				
NDRT Form H	Block 1				
	Income	0.32	0.31	.10	1.03
	VCI	0.13	0.03	.38	4.07**
	$R^2 = .16, F(2, 101) = 9.91**$				
	Block 2				
	Income	0.07	0.31	.02	0.22
	VCI	0.08	0.03	.24	2.37*
	Broad Reading	0.17	0.06	.32	3.05**
	$\Delta R^2 = .07, \Delta F(1, 100) = 9.32**$				

Note: NDRT = *Nelson–Denny Reading Test*; VCI = Verbal Comprehension Index; WAIS-IV = *Wechsler Adult Intelligence Scale–Fourth Edition*; WJ-III-Ach = *Woodcock–Johnson Tests of Achievement–Third Edition*.
*p < .05. **p < .01.

Table 3. Correlations Between NDRT Literal and Interpretive Items, WAIS-IV Scores, and WJ-III-Ach Broad Reading.

	Form G		Form H	
	Literal	Interpretive	Literal	Interpretive
WAIS-IV				
VCI	.37**	.42**	.25**	.42**
PRI	.10	.14	−.03	.25**
WMI	.28**	.22*	.14	.32**
PSI	.22*	.20*	.29**	.13
FSIQ	.26**	.36**	.21*	.40**
WJ-III-Ach				
Broad Reading	.35**	.39**	.35**	.41**

Note: FSIQ = Full-Scale Intelligence Quotient; NDRT = *Nelson–Denny Reading Test*; PRI = Perceptual Reasoning Index; PSI = Processing Speed Index; VCI = Verbal Comprehension Index; WAIS-IV = *Wechsler Adult Intelligence Scale–Fourth Edition*; WJ-III-Ach = *Woodcock–Johnson Tests of Achievement–Third Edition*; WMI = Working Memory Index.
*p < .05. **p < .01.

VCI, relative to other indices of the WAIS-IV. The VCI is composed of three subtests that measure overlapping and distinct aspects of verbal comprehension. All three subtests tap crystallized knowledge, long-term memory, and verbal expression (Wechsler et al., 2008). Vocabulary also assesses language development and verbal concept formation. Similarities measures associative, categorical, and abstract reasoning. Information assesses general factual knowledge

of information. In the current data, these subtests were correlated .28 to .49.

WJ-Ach-III also was consistently and modestly associated with NDRT scores. Subscales of the Broad Reading index are Letter–Word Identification, Reading Fluency, and Passage Comprehension (Mather & Woodcock, 2001). Letter–Word Identification assesses word identification skills. Reading Fluency measures the speed at which simple sentences can be read and comprehended. Passage Comprehension measures reading comprehension for sentences and short reading passages. In the current data, these subtests were correlated .33 to .37. Correlations between the VCI and Broad Reading subtests ranged from −.01 to .58 (*Mdn* = .33).

Since the VCI and Broad Reading were consistently associated with passageless scores on the NDRT Forms, follow-up analyses were conducted to determine if VCI and Broad Reading subscales were differentially associated with NDRT scores (Table 5). Of the VCI subscales, Vocabulary and Information were most consistently and strongly associated with NDRT scores. For the Broad Reading component scores, Letter–Word Identification and Passage Comprehension had the most consistently significant correlations with NDRT scores.

Discussion

Reading comprehension items from the NDRT can be answered correctly without reading the associated passage

Table 4. Correlations Between NDRT Items by Content Type, WAIS-IV Scores, and WJ-III-Ach Broad Reading.

	Form G			Form H		
	Humanities	Social Science	Science	Humanities	Social Science	Science
WAIS-IV						
VCI	.39**	.35**	.32**	.21*	.31**	.34**
PRI	.23*	.08	.03	.18	.06	.10
WMI	.19*	.16	.23*	.37**	.02	.21*
PSI	.10	.19*	.28**	.03	.12	.29*
FSIQ	.34**	.30**	.30**	.26**	.20*	.33**
WJ-III-Ach						
Broad Reading	.29**	.24**	.43**	.28**	.22*	.43**

Note: FSIQ = Full-Scale Intelligence Quotient; NDRT = *Nelson–Denny Reading Test*; PRI = Perceptual Reasoning Index; PSI = Processing Speed Index; VCI = Verbal Comprehension Index; WAIS-IV = *Wechsler Adult Intelligence Scale–Fourth Edition*; WJ-III-Ach = *Woodcock–Johnson Tests of Achievement–Third Edition*; WMI = Working Memory Index.
p < .05. **p* < .01.

Table 5. Correlations Between NDRT Scores and WAIS-IV VCI Subtests and WJ-III-Ach Broad Reading Subtests.

	NDRT	NDRT	Form G		Form H		Form G			Form H		
	G	H	Literal	Interp	Literal	Interp	Hum	Soc Sci	Sci	Hum	Soc Sci	Sci
VCI Subscales												
Vocabulary	.33**	.41**	.28**	.27**	.29**	.37**	.33**	.20*	.20*	.19*	.35**	.29**
Similarities	.29**	.09	.20*	.25**	.03	.11	.21*	.25**	.16	−.06	.16	.08
Information	.46**	.46**	.34**	.41**	.30**	.45**	.36**	.32**	.33**	.31**	.21*	.41**
Broad Reading subscales												
Letter–Word Identification	.40**	.36**	.27**	.36**	.24**	.33**	.28**	.22*	.34**	.25**	.19*	.29**
Reading Fluency	.34**	.32**	.28**	.25**	.29**	.25**	.13	.22*	.34**	.16	.15	.32**
Passage Comprehension	.34**	.39**	.28**	.29**	.24*	.39**	.28**	.15	.31**	.24**	.21*	.34**

Note: Hum = Humanities; Interp = Interpretive; NDRT = *Nelson–Denny Reading Test*; Sci = Science; Soc Sci = Social Science; VCI = Verbal Comprehension Index; WAIS-IV = *Wechsler Adult Intelligence Scale–Fourth Edition*; WJ-III-Ach = *Woodcock–Johnson Tests of Achievement–Third Edition*.
p < .05. **p* < .01.

and passageless response rates are substantially greater than chance. Passageless administration scores pose a major threat to the validity of the NDRT. If reading comprehension items can be answered without reading the associated passage, the test fails in its core mission to assess reading comprehension.

Consistent with predictions, passageless NDRT scores were significantly associated with verbal comprehension and reading skills. These findings suggest that when the NDRT is administered under standardized conditions, individual differences in IQ and reading skills may be significantly associated with reading comprehension scores. Individual differences in IQ and reading skills may introduce significant construct-irrelevant variance into performance scores on the reading comprehension component of the NDRT. Subsequently, the NDRT could be a less valid

measure of reading comprehension for persons with relatively stronger vocabularies, greater funds of knowledge, and superior broad reading skills. For example, false negatives on the NDRT reading comprehension test could be more likely for persons with better academic and intellectual resources than those without because persons with greater resources answered more questions correctly without doing the associated reading.

To illustrate, in the current data, persons who had a VCI score greater than 120 (e.g., greater than the 91st percentile for the WAIS-IV age-corrected normative sample), answered 18 to 19 items correctly on passageless administration of the NDRT. These are response rates of 47% to 50% correct. In contrast, persons with a VCI score of 100 or less (e.g., at the 50th percentile or below for the WAIS-IV age-corrected normative sample) answered only 13 to 14 items correctly in

passageless administration, response rates of 34% to 37%. Clearly, verbal comprehension skills may have a clinically significant impact on passageless reading comprehension scores. In future studies, to fully test the hypothesis that the NDRT has differential validity based on IQ and reading skills, the NDRT would have to be administered under standard administration conditions.

If future research does support the hypothesis that the NDRT has differential validity based on IQ and achievement scores, this finding would be troubling. That is, ideally, a reading comprehension test will assess knowledge obtained from reading a passage and not knowledge from extraneous factors such as general information knowledge. To the extent that general information knowledge influences reading comprehension test scores, construct-irrelevant variance is introduced to the test score and the validity of that test score is negatively affected. As a consequence, reading comprehension deficits may be more difficult to detect with the NDRT in high-achieving and bright students. An alternative perspective on our data is that the NDRT will be a less valid test of reading comprehension in persons with weak reading skills because these persons will be more likely to rely on previous knowledge in an attempt to answer reading comprehension test items. Future studies to test these competing hypotheses are needed.

Converging Evidence on the NDRT

Several findings from Coleman et al. (2010) were replicated in the current study. Accurate response rates averaged 42% to 45%, and Coleman et al. reported nearly identical rates of 44% to 47%. In both studies, science questions were answered correctly more often than humanities items. Given that reading passages were extracted from existing texts, the fact that science questions are answered more easily than humanities is interesting. Perhaps there is less diversity in science curriculum and texts in high school and college than for the humanities, thus there may be more uniform and widespread exposure to science than humanities content across test takers. Finally, in both studies, participants with a potential LD answered passageless NDRT items with about the same success as participants who did not have an LD. Thus, college students, regardless of LD status, may have accumulated sufficient general information knowledge in their educational careers to answer passageless reading comprehension test items with comparable accuracy.

Another critical finding for which there is converging evidence is that the NDRT forms are not parallel. Form G has an imbalance as to correct response rates for passageless items. Similar to Coleman et al. (2010), in our data, interpretive items were answered correctly more often than literal items. For Form H, literal and interpretive items were more balanced in their response rates for passageless administration. Practitioners may prefer to use Form H because of the more

balanced item composition, especially if they attend to response rates for literal versus interpretive items.

Building a Better Reading Comprehension Test

Passageless response rates are problematic for the NDRT and GORT-3, two of the most popular tools to assess reading comprehension. The NDRT suffers from other problems, such as outdated normative data. A major revision of the NDRT is needed, at minimum. This effort should involve careful development and testing of passage comprehension items to ensure they cannot be answered correctly during passageless administration. Alternately, perhaps a fresh approach to assessing reading comprehension in adults is in order. A first and reasonable solution to the vexing problem of passageless response rates would be to use reading passages to which no respondent has had previous exposure. Fiction passages, in a variety of formats (e.g., historical fiction, science fiction, drama), may be a viable alternative to using existing texts. Creative writing programs could pair with test developers to create passages that varied in length, complexity, vocabulary, and sentence structure. Next, test items could be written for these passages, and then these items could be systematically assessed for passage independence prior to publication of the instrument. It is clear that passage independence of items should not be assumed but should be rigorously tested.

Limitations

The current study offers new data that link individual difference factors with passageless administration scores on the NDRT, but there are limitations. Some participants reported previous exposure to the WAIS-III or WAIS-IV during testing. It would have been wise to systemically assess for prior exposure to the WAIS-IV, previous editions of the WAIS, the WJ-III-Ach, and the NDRT and to determine how prior exposure may have affected findings, but unfortunately these data were not collected. Furthermore, the sample was one of convenience, and past diagnoses of LD and ADHD were not verified. Another limitation is that multiple analyses were conducted on the data and there were no corrections for multiple comparisons. However, effect sizes of associations were of interest, in addition to significance testing, and the magnitude and patterns of linkages among the VCI, Broad Reading, and NDRT scores were robust and thus are likely to be clinically meaningful rather than spurious findings.

Declaration of Conflicting Interests

Funding

References

American College Health Association. (2011). *National College Health Assessment II: Fall 2011 reference group executive summary*. Hanover, MD: Author.

Brown, J. I., Fishco, V. V., & Hanna, G. S. (1993). *Nelson–Denny Reading Test*. Itasca, IL: Riverside.

Coleman, C., Lindstrom, J., Nelson, J., Lindstrom, W., & Gregg, K. N. (2010). Passageless comprehension on the Nelson–Denny Reading Test: Well above chance for university students. *Journal of Learning Disabilities*, *43*, 244–249.

Educational Testing Service, Office of Disability Policy. (2007). *Policy statement for documentation of a learning disability in adolescents and adults* (2nd ed.). Princeton, NJ: Author.

Educational Testing Service, Policy Information Center. (2007). *ETS policy notes: Students with learning disabilities transitioning from high school to college*. Princeton, NJ: Author.

Katz, L. J., Goldstein, G., & Beers, S. S. (2001). *Learning disabilities in older adolescents and adults*. New York, NY: Kluwer Academic/Plenum.

Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the Gray Oral Reading Test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading*, *10*, 363–380.

Law School Admission Council. (2012). *Guidelines for documentation of cognitive impairments*. Newtown, PA: Author.

Mather, N., & Woodcock, R. W. (2001). *Woodcock–Johnson Tests of Achievement–Third Edition, Examiner's manual*. Itasca, IL: Riverside.

Patterson, M. B. (2008). Learning disability prevalence and adult education program characteristics. *Learning Disabilities Research & Practice*, *23*, 50–59.

Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics*, *104*, 1351–1359.

Stetson, R., Stetson, E. G., & Sattler, J. M. (2001). Assessment of academic achievement. In J. M. Sattler (Ed.), *Assessment of children: Cognitive applications* (4th ed., pp. 576–609). San Diego, CA: J. M. Sattler.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*, 623–628.

Wechsler, D., Coalson, D. L., & Raiford, S. E. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV), technical and interpretive manual*. San Antonio, TX: Pearson.