Evaluation of the Psychometric Properties of an Online Evaluation Tool for Dyslexia

David P Hurford*, Lauren P Decker and Samantha P Conway

Center for READing, Pittsburg, Pittsburg State University, Broadway, Pittsburg

***Corresponding author:** David P. Hurford, Center for READing, Pittsburg State University, Pittsburg, Kansas, United States, Tel: +1 6202354534;
E-mail: dphurford@pittstate.edu

Abstract

The purpose of the study was to examine the psychometric properties of an online dyslexia evaluation tool (ODET). The ODET is comprised of 14 subtasks that evaluate phonological processing (Word Deletion, Syllable Deletion, Onset-Rime Deletion, Phoneme Deletion, Onset-Rime Blending, Phoneme Blending, Segmentation and Dynamic Segmentation), letter knowledge (Letter Knowledge and Letter Sound Knowledge) and reading (Word Knowledge, Word Building Knowledge, Spelling and Sentence Completion). The ODET provides percentage correct, standard scores and At-Risk scores for each subtest and a Global At-Risk Score that provides information whether the child is At-Risk for dyslexia. To examine the psychometric properties, 40 children (16 females and 24 males) were evaluated with the ODET, the CTOPP-2, the WIAT-III, and the PPA. Results indicated that the ODET had appropriate psychometric properties.

Keywords: Reading difficulties; Dyslexia; Evaluation; Online

Introduction

Dyslexia refers to a reading disability that is neurobiologically based, occurs as a function of poor phonological processing skills, is characterized by poor fluency and spelling, is genetically based and lasts a lifetime. Individuals with dyslexia have significant difficulties manipulating the sounds of language, mapping graphemes onto phonemes, decoding graphemes into phonemes, and synthesizing the decoded phonemes into syllables and words. These difficulties affect fluent reading and diminish comprehension, are associated with spelling and writing difficulties, and a host of other related difficulties. Individuals with dyslexia can become competent readers if the cause of their reading failure is identified and explicit instruction is provided [1-12].

Dyslexia is the most common learning difference and its prevalence has been estimated to be 15% to 20%. In addition, approximately 34% of children in the United States are not reading at the Basic Level, which represents a rudimentary ability to read [13]. For individuals with dyslexia, reading acquisition can be extremely challenging [14]. Appropriate intervention and remediation efforts can result in competent reading [15]. However, prior to access to interventions, appropriate evaluations must occur that accurately identifies students with dyslexia. Unfortunately, many students with dyslexia are not diagnosed until they have experienced reading failure for far too long, in some cases years [16]. Highly effective interventions for dyslexia exist [17-19], but far too many students are not identified as early as is scientifically possible. As a result, these children continue to struggle unnecessarily. Early identification is essential [20-23].

The ramifications of dyslexia often extend far beyond the classroom. Children with dyslexia not only experience academic difficulties but are also more likely to have poor self-esteem, anxiety, and Post Traumatic Stress Disorder [24-28]. These children are more likely to abuse substances, become depressed, have suicidal ideation and and commit suicide [29,30]. Children with dyslexia are more likely to experience parental physical abuse[31], are at a greater risk of not graduating from high school, to be adjudicated as juveniles, and later as adults[31-34].

Fortunately, the above-mentioned difficulties can be prevented with appropriate interventions if the individuals can be identified prior to or early in their academic careers [35-38]. The following measures have been used to accurately predict reading failure: letter name and sound letter knowledge, phonemic synthesis, phonemic analysis, word reading, non-word reading, comprehension and spelling.

The Online Dyslexia Evaluation Tool (ODET) was developed to be an evaluation tool that could be used to examine the phonological processing, reading, spelling and comprehension skills listed above as a preliminary step in the diagnosis of dyslexia and reading difficulties. The present study sought to examine the ODET's psychometric properties.

The ODET has been used mostly as a research tool and on a

small scale by parents and teachers for the past 20 years. Prior to presenting it to a wider user base, it was necessary to examine its psychometric properties in more detail. The validity study was undertaken to examine the psychometric properties of the ODET to ensure that it was a valid measure of reading and its subcomponents.

The ODET is comprised of 14 subtasks that measure phonological processing (Word Deletion, Syllable Deletion, Onset-Rime Deletion, Phoneme Deletion, Onset-Rime Blending, Phoneme Blending, Phoneme Segmentation and Dynamic Segmentation), knowledge of letters (Letter Knowledge and Letter Sound Knowledge), reading (Word Knowledge and Word Building Knowledge), comprehension (Sentence Completion) and spelling (Spelling) in the order just listed. The ODET was formatted to be web based and allows parents, teachers and other professionals to assess their children, students and clients on the tasks listed above.

The ODET provides interpretation of the findings in an eight-page Summary of the Assessment report that is generated by an algorithm after the ODET has been administered. The report provides demographic information concerning the child (name, date of birth, date of examination, age and gender), explanation of the nature of standard scores, descriptions of phonemic analysis, phonemic synthesis, letter knowledge, reading ability, spelling ability and comprehension along with specific results on each of the subtests. A Global At-Risk Score, its standard score and a description of the meaning of those values is presented which is the main metric that indicates whether the child performed in the risk for reading difficulties range. Lastly, an interpretive summary is provided that lists the areas of weaknesses and their meaning so that Summary of Assessment can be used by school personnel to determine if further assessment will be required or to develop appropriate interventions, if necessary.

Material and Methods

Participants

Forty students enrolled in Midwest elementary schools participated in the present study see **Table 1**. All of the students' families were economically in the lower middle to upper middle-class range. A *t*-test was performed examining the mean age of the male and female participants which indicated that there was no effect of gender, $t(38)=0.87$, $p>.39$. Although there were more males, the results of a Chi-Square analysis indicated that there were no significant differences in the number of males and females between groups, $\chi2 (1)=1.60$, $p>.21$.

Table 1 Demographic Information of the Participants

Characteristic	Gender	
	F	**M**
Age	9.25	9.02
Ethnicity		
White	12	19
Hispanic	3	4

African-American	1	1
Parental Education Level		
High School Graduate	6	11
Some College	3	4
College Graduate	3	9

Tasks and materials

Online Dyslexia Evaluation Tool (ODET): The ODET was standardized with 2,240 students ranging in age from 5 years to 12 years prior to this study. The number of items for each subtask, and the ability assessed, along with the requirements for performance can be seen in **Table 2**. During administration, the ODET automatically tracks passed and failed items and discontinues a task if the child misses five consecutive items. One of the advantages of the ODET is that the child is not aware that he or she met the discontinuation rule.

Table 2 Number of Items, Ability Assessed, and Task Requirements by Task on the ODET.

Task	Items	Ability Assessed	Task Requirements
Word Deletion	**16**	Phonological Processing	Deleting a word from compound word
Syllable Deletion	**16**	Phonological Processing	Deleting a syllable from a word.
Onset-Rime Deletion	**16**	Phonological Processing	Deleting the onset/rime from a word
Phoneme Deletion	**16**	Phonological Processing	Deleting a phoneme from a word.
Onset-Rime Blending	**16**	Phonological Processing	Blending onsets and rimes to form words.
Phoneme Blending	**16**	Phonological Processing	Blending phonemes to form words.
Segmentation	**16**	Phonological Processing	Segmenting words into phonemes
Dynamic Segmentation	8	Phonological Processing	Guided segmenting words into phonemes
Letter Knowledge	**52**	Letter Names	Naming specific letters.
Letter-Sound Knowledge	26	Letter Sounds	Indicating the sounds that correspond to letters.
Word Knowledge	50	Word Reading	Reading words.
Word Building	40	Reading Non-words	Reading non-words.
Sentence Completion	25	Comprehension	Indicating the missing word in a sentence.
Spelling	15	Spelling	Spelling words.

> Note: All of the items in the subtasks are scored as correct (1) or incorrect (0), except for Spelling in which a correctly spelled word is scored as 2, a phonetically-correctly spelled word as 1, and an incorrectly spelled word as 0.Total correct for Spelling is therefore 30.

The ODET provides percentage correct, standard scores and descriptors for each subtask. It was found that the Complex Computation Methods (CCM) provided not only the best accuracy, but the lowest false negative rate compared to other mechanisms used for predicting reading failure [39]. The CCM utilized a weighting system based on a particular variable's ability to predict poor reading performance. The weighting system was developed from the canonical structure coefficients examining the ability of this method to predict reader group membership. The specific weights were created by multiplying the canonical structure coefficients for each variable by 10 and rounding to whole numbers. For example, the canonical structure coefficient for Letter Knowledge was .917682. After being multiplied by 10 and rounded, the weight for Letter Knowledge was 9 (.917682 x 10=9.17682 ≈ 9). Each subtask had a different weight based on its canonical structure coefficient. In addition, the weight for each sutask was further differentiated based on performance. A raw score that was 1 sd or more below the mean was assigned the full At-Risk weight, a raw score that was between -1 *sd* and -0.5 *sd* was assigned 80% of the weight and a raw score between -0.5 and the mean was assigned 60% of the weight. A raw score equal to or above the mean did not receive an At-Risk weighting. For example, the At-Risk weightings for the Letter Knowledge subtask were 9, 7 (9 x .80=7.2 ≈ 7) and 5 (9 x .60=5.4 ≈ 5). The Global At-Risk Score (GARS) was determined by summing all of the At-Risk weightings from each subtask. Poor performance on one or several of the subtasks would result in larger GARS values. A GARS above 1 above the mean GARS resulted in a determination of "At-Risk." Using the above procedure resulted in highly accurate identification of students either At-Risk or experiencing reading failure.

The time to administer the ODET varied depending on the participant's abilities, but ranged from 40 to 60 minutes. Administration of the ODET typically begins with instructions regarding the nature of the test, the importance of accurately administrating the test and recording the responses and that any of the tasks can be

Figure 1 Example of the instructions for the ODET.

Figure 2 Example of the practice trials for the Word Deletion task.

paused if there was a need. However, the trained staff was trained prior to administration regarding these issues. These instructions were contained within each task **Figure 1**. The tasks are presented in such a fashion that the administrator simply presents to the child exactly what is provided on the computer's screen and then records the child's response as correct or incorrect by clicking on a radio button. Examples of correct and incorrect performance are listed for the administrator to view **Figure 2**.

Comprehensive Test of Phonological Processing 2 (CTOPP-2)

The CTOPP-2 was used to evaluate the validity of the ODET's phonological processing subtasks [40]. The CTOPP-2 consists of three subtests that assess phonological awareness (Elision, Blending Words and Phoneme Isolation), two subtests that assess phono-

logical memory (Memory for Digits and Nonword Repetition) and two subtests that assess rapid naming (Rapid Digit Naming and Rapid Letter Naming). The CTOPP-2 also includes two supplemental subtests (Blending Non-words and Segmenting Non-words). For the purpose of the present study, the Elision, Blending Words, Blending Non-words and Segmenting Non-words subtests were used as they related more closely to providing evidence for validity for the phonological processing subtasks of the ODET than the others subtests (see the top panel of **Table 3** for an explanation of the CTOPP-2 subtests). The ODET does not assess phonological memory or rapid naming. The Phoneme Isolation subtest of the CTOPP-2 has no analogous subtask on the ODET. The CTOPP-2 has been extensively evaluated in research and in practice and has been established in terms of its psychometric properties [41].

Table 3 Number of Items, Ability Assessed and Task Requirements by Subtest on the CTOPP-2, WIAT-III and PPA.

Subtest	Items	Ability Assessed	Task Requirements
Elision	34	Phonological Processing	Deleting various segments from words
Blending Words	33	Phonological Processing	Blending sounds into words
Blending Nonwords	31	Phonological Processing	Blending sounds into nonwords
Segmenting Nonwords	31	Phonological Processing	Segmenting nonwords into their sounds
WIAT-III Subtests			
Reading Comprehension	84	Reading Comprehension	Reading comprehension skills
Word Reading	75	Word Reading	Accuracy of reading individual words
		Word Reading Speed	Speed of reading individual words
Spelling	63	Spelling	Spelling dictated words
PPA Subtests			
Print Knowledge	17	Identifying aspects of print	Identify that a book is for reading, identify capital and lower case letters, first and last words, vowels, etc.
Sound-Symbol	11	Letter-sounds knowledge	Recognize various consonants either at the beginning or end of words and identify the name for that letter

Wechsler Individual Achievement Test-III (WIAT-III)

The WIAT-III has 16 subtests that evaluate listening speaking, writing, and mathematical reasoning. Of the 16 subtests, only three subtests were relevant to examining the validity of the

ODET: Word Reading, Reading Comprehension and Spelling (see the middle panel of **Table 3** for an explanation of these WIAT-III subtests) used in this study. The Word Reading subtest evaluates the accuracy and speed of reading individual words. As a result,

4

this subtest produces two scores, one based on accuracy and one based on speed. Reading speed is determined by evaluating the number of correctly read words 30 seconds into the task. The WIAT-III has adequate psychometric properties [42,43].

Phonological and Print Awareness Scale (PPA).

The PPA is comprised of six subtests. Four of the subtests assess phonological awareness (Rhyming, Initial Sound Matching, Final Sound Matching and Phonemic Awareness). The other two subtests evaluate print awareness (Print Knowledge and Sound Symbol). It was these two latter subtests that were used in the present study to examine the ODET's Letter Knowledge and Letter Sound Knowledge subtests (see the bottom panel of **Table 3** for an explanation of these PPA subtests). The CTOPP-2 was already being utilized to evaluate the phonological processing subtasks of the ODET. The PPA has adequate psychometric properties for the two subtests included in this study [44,45].

Procedure

Each participant was assessed with the ODET, CTOPP-2, WIAT-III and PPA individually in a quiet and vacant room in his or her school across three different sessions that were counterbalanced in order of presentation. The participants were also assessed in accordance with the standardization procedures outlined for each test.

Results and Discussion

Reliability

Reliability was evaluated in several ways. First, Cronbach's Alpha, which is a measure of internal reliability, ranged from .96 to .99 for the various subtasks. As can be seen in **Table 4**, test-retest and inter-rater reliability ranged from .96 to .99. To derive these values the Expectation Maximization algorithm [46] was used as there were several tasks that had very severe restrictions of range that resulted in correlation coefficients that did not adequately reflect the nature of the relationship. As an example, the participants all scored between 50 and 52 on the Letter Knowledge task (maximum score=52), which is a task that examined the participants' knowledge of letter names, at Time 1 and Time 2. This restriction of range resulted in a correlation coefficient of .11 as the relationship between Time 1 and Time 2 was at ceiling. Employing the EM algorithm resulted in more accurate reliability estimates.

Table 4 Test-Retest and Interrater Reliability.

	Type of Reliability	
ODET Subtask	Test-Retest	Interrater
Letter Knowledge	.97	.96
Letter-Sound Knowledge	.98	.96
Word Knowledge	.99	.99
Word Building Knowledge	.98	.99
Spelling	.96	.98
Sentence Completion	.97	.97
Deletion	.97	.97
Blending	.95	.98
Phonemic Segmentation Deletion	.97	.99
Dynamic Segmentation	.98	.98
Prompts (Dynamic Segmentation)	.97	.98

Validity

Face validity: Face validity refers to the notion that the test appears to measure the intended purpose of the test and since the ODET measures phonological processing, letter knowledge, reading, spelling and comprehension, which are areas in which individuals struggle who have dyslexia/reading difficulties, the ODET appears to have face validity. In addition, a panel of school psychologists (*n=6*, mean number of years as practicing school psychologists=14.3 years) who were experts in evaluating children's phonological processing, reading, spelling and comprehension skills were provided with the content of the ODET and their unanimous opinion was that the ODET had face validity.

Content validity: Content validity evaluates the content of the test to determine whether the test appropriately samples the domain of interest, in this case phonological processing, letter knowledge, reading, spelling and comprehension. An enormous literature exists that has examined these issues for decades and the content of the ODET is similar to those tasks that have been used in that literature. In fact, the ODET grew out of that literature, particularly as it relates to the identification of poor reading skills [47]. Nearly all instruments designed to evaluate the possibility of reading difficulties include measures that evaluate phonological processing, knowledge of letters including their names and associated sounds, reading of words and non-words, spelling and comprehension (e.g., Woodcock Reading Mastery Tests, 3rd Edition; Comprehensive Test of Phonological Processing, 2nd Edition, etc.). The panel of school psychologists described above was asked to comment on the content of the ODET and their collective opinions were in agreement that the ODET had sufficient content validity.

Criterion related validity: Criterion related validity involves examining how well a test is related to a particular criterion or outcome. In this regard, the ODET's results should be related to various aspects of reading including phonological processing, letter knowledge, reading, spelling and comprehension, as these are the components evaluated with the ODET. Typically, criterion related validity is established by examining the relationship between a particular test and other tests that purport to measure the same criterion. In this case, reading and its subcomponents.

To examine criterion related validity, the subtasks of the ODET were compared to three other tests and their subtasks that measure the same domain. The CTOPP-2 was used to examine the ODET's measures of phonological processing, the WIAT-III was used to examine the ODET's measures of reading, spelling, comprehension, and the PPA was used to examine the ODET's measures of letter knowledge (Letter Name Knowledge and Letter Sound Knowledge).

The four CTOPP-2 subtests that measured the same phonological processing content areas as the ODET were moderately to strongly corrected (see the top panel of **Table 5**). The Deletion score from the ODET is comprised of four subtasks (Word Deletion, Syllable Deletion, Onset-Rime Deletion and Phoneme Deletion) which are similar to the Elision subset from the CTOPP-2 and these measures are significantly correlated (.58). The Blending score from the ODET is comprised of the Onset-Rime and Phoneme Blending tasks that were compared to the Blending Words and Blending Non-words subtests of the CTOPP-2. As can be seen in **Table 5**, these subtests were significantly correlated at .54 and .58, respectively. The Phoneme Segmentation, Dynamic Segmentation and Prompts tasks form the ODET were significantly correlated with their analogous subtests from the CTOPP-2; .35, .56 and -.56, respectively. The difference between the Phonemic Segmentation task from the ODET and the Segmenting Non-words from the CTOPP-2 is that the ODET version uses real words and the CTOPP-2 version uses non-words. The Dynamic Segmentation task has half the number of items as the Phonemic Segmentation (8 vs. 16, respectively), and requires prompts to be administered if the child fails an item. The prompts help to determine what

scaffolding is necessary for the child to successfully perform on segmenting skills. This provides informative information regarding the child's capability. There are seven levels of prompts that successively provide clues and suggestions regarding how to engage in the segmentation task. For example, a Level 1 prompt simply emphasizes each sound in the word for the child and then the child is asked to segment the word into its sounds. A Level 7 prompt asks the child to listen as the administrator models segmentation and then the child is asked to segment the words into its sounds. Like Vygotsky's Zone of Proximal Development, the level of prompt probes the child to determine what level of scaffolding is necessary for the child to produce the correct answer. As a result, the Dynamic Segmentation task produces two scores, the number correct and the number and level of prompts. As the level of prompt increases from Level 1 to Level 7, the poorer the performance and the greater the need for scaffolding. Therefore, the level of prompt will be negatively correlated with other measures. As can be seen in **Table 5**, both scores are significantly correlated with the CTOPP-2's Segmenting Non-words subtests.

Table 5 Correlation Coefficients of the ODET, CTOPP-2, WIAT-III, and PPA Subtests.

CTOPP-2 Subtests				
ODET Subtasks	**Elision**	**Blending Words**	**Blending Nonwords**	**Segmenting Nonwords**
Deletion	.58**	.40**	.345*	.55**
Blending	0.29	**.54****	**.58****	.50**
Phoneme Segmentation	0.26	0.17	.36*	**.35***
Dynamic Segmentation	.38**	.448**	.57**	**.56****
Prompts	-.36**	-.48**	-.61***	**-.56****
WIAT-III Subtests				
ODET Subtasks	**Word Reading**	**Word Reading Speed**	**Reading Comprehension**	**Spelling**
Word Knowledge	**.65****	**.76****	.83***	.50**
Word Building	**.57****	**.62****	.56***	.52**
Sentence Completion	.50**	.75***	**.80****	.44*
Spelling	.76***	.79***	.73***	**.78****
PPA Subtests				
ODET Subtasks	**Print Knowledge**		**Sound-Symbol**	
Letter Knowledge	**.40***		.63***	
Letter-Sound Knowledge	0.17		**.81***	

Note: *= p<.03, **=p<.01, ***=p < 001. Bolded values indicate subtasks and subtests purported to measure same content. Poor performance on Prompts is indicated by larger scores. As a result, it was negatively associated with the other measures.

The WIAT-III subtests that measured the same reading, comprehension and spelling subtasks as the ODET were significantly (correlated see the middle panel of **Table 5**). The Word Knowledge and Word Reading tasks both measure the ability of the individual to read words. Word Reading Speed was also highly correlated with Word Knowledge. These two subtests of the WIAT-III were also similarly correlated with the

Word Building subtask of the ODET. The Word Building subtask asks the individual to read nonwords. The Sentence Completion task from the ODET is a cloze task in which the individual is required to indicate the missing word in a sentence. This subtask was thought to be a proxy of comprehension, but it was very strongly correlated with the Reading Comprehension subtest of the WIAT-III, indicating that it is more of a viable measure of

reading comprehension than was originally thought. The Spelling subtask from the ODET and the Spelling subtest of the WIAT-III were strongly correlated.

The PPA subtests that measured the same letter knowledge and letter sound knowledge subtasks as the ODET were significantly correlated see the bottom panel of **Table 5**. The Letter Knowledge subtask of the ODET measures the individual's knowledge of letter names and is significantly correlated with the PPA subtest of Print Knowledge. In addition to measuring letter name knowledge, the Print Knowledge subtest also measures knowledge of print, such as concepts regarding books, letters, first and last words in sentences, punctuation (e.g., point to the sentence with the quotation marks), vowels and capitalization. Although 10 out of the 17 items on Print Knowledge directly assess letter name knowledge, all of the items on the Letter Knowledge subtask of the ODET does. Although the correlation was significant between

the two subtasks, it likely would have been much stronger if both directly and exclusively measured letter name knowledge. This hypothesis is supported by the strong correlation between the Letter Knowledge and Sound Symbol subtasks in which Sound Symbol only includes stimuli that are letters. The Letter Sound Knowledge subtask of the ODET and the Sound Symbol subtest of the PPA were very strongly correlated (.81).

Predictive validity: Eleven participants in the validity study scored below the 25th percentile, which also corresponded to a score that was greater 14 on the GARS from the ODET. Those participants were considered to be "At-Risk" for reading failure as the ODET was designed to identify children At-Risk. The data from these children were evaluated as above to evaluate predictive validity. As can be seen in **Table 6**, the correlation coefficients were even stronger for the At-Risk group than they were for the entire group of participants.

Table 6 Correlation Coefficients of the ODET, CTOPP-2, WIAT-III, and PPA Subtests for Participants Identified as At-Risk for Reading Failure.

CTOPP-2 Subtests				
ODET Subtasks	**Elision**	**Blending Words**	**Blending Nonwords**	**Segmenting Nonwords**
Deletion	**.60**[*]	.54[*]	.46	.55
Blending	-.02	**.80**[**]	**.75**[**]	.57[**]
Phoneme Segmentation	-.06	.07	.25	**.26**
Dynamic Segmentation	.05	.39	.77[**]	**.64**[*]
Prompts	-.21	-.57	-.79[**]	**-.69**[*]
WIAT-III Subtests				
ODET Subtasks	**Word Reading**	**Word Reading Speed**	**Reading Comprehension**	**Spelling**
Word Knowledge	**.85**[***]	**.90**[**]	.96[***]	.74[**]
Word Building	**.82**[***]	**.66**[*]	.61[*]	.81[**]
Sentence Completion	.87[**]	.93[***]	**.95**	.75
Spelling	.96[***]	.93[***]	.90[***]	.87[***]
PPA Subtests				
ODET Subtasks	**Print Knowledge**	**Sound-Symbol**		
Letter Knowledge	**.50**[*]	.71[**]		
Letter-Sound Knowledge	0.17	**.88**[**]		

Note: [*]= p<.03, [**]=p<.01, [***]=p < 001. Bolded values indicate subtasks and subtests purported to measure same content. Poor performance on Prompts is indicated by larger scores. As a result, it was negatively associated with the other measures.

Construct validity: Construct validity examines how well the test measures the underlying theoretical constructs of a domain. In this case, the domain concerns reading and its subcomponents: phonologic processing, letter knowledge, word and nonword reading, spelling and comprehension. Since all of these measures assess reading, all of the subtasks should be related. The subtasks should also be highly related as they measure the same component of reading. Reading acquisition abilities and its subskills increases with age until reading becomes proficient. As a result, the measures on the ODET should be affected by age. Lastly, the ODET should be able to differentiate between good

and poor readers as it was designed to identify children who were At-Risk for reading failure or who were experiencing reading.

Construct validity was evaluated by examining the correlation coefficients between and among the subtasks of the ODET, by performing an exploratory factor analysis, and by examining the performance of the participants by age and by examining whether the ODET could differentiate between good and poor readers. As can be seen in Table 7, of the 136 correlation coefficients, all but 14 are significant. Since the ODET evaluates components of reading, it was expected that nearly all of the correlation coefficients would be statistically significant. It was also expected that the variables

that measured similar subdomains would be highly significant and the results confirm this expectation; Letter Knowledge and Letter Sound Knowledge (.73) and Word Knowledge, Word Building Knowledge, Spelling and Sentence Completion (.70, .82 and .81, respectively). The phonological processing measures were also significantly correlated. Interestingly, Phoneme Deletion had the largest correlation coefficient with Deletion, which is the composite variable comprised of Word Deletion, Syllable Deletion, Onset-Rime Deletion and Phoneme Deletion (.83) and Phoneme Blending had the largest correlation coefficient with Blending, which is the composite variable comprised of Onset-Rime Blending and Phoneme Blending (.95). The ability to process and manipulate phonemes is an important skill as Phoneme Deletion and Phoneme Blending were significantly and strongly related to all of the other variables. The literature has demonstrated the same relationship for decades.

The variable that most strongly correlated with the other variables was the GARS. As described above, the GARS is a composite score comprised of the sum of all of the At-Risk weighted subtask scores. A weighted At-Risk value is associated with each task based on the individual's performance. A task's weight was derived by determining the importance of that task in identifying individuals who are At-Risk for reading failure (Blinded for Review, 2002). The more important that task was in differentiating between good and poor readers and the poorer the performance on that task, the greater the At-Risk score. As can be seen in **Table 7**, the GARS was highly correlated with all of the variables ranging from -.34 to -.82 (*M*=-.69). The reason that the GARS was negatively correlated with the other values was that it was inversely related to performance. The higher the GARS score which is the At-Risk indicator, the poorer the performance.

Table 7 Correlation coefficients by ODET subtask.

Subtask	LK	LSK	WK	WBK	SC	DEL	WD	SD	ORD	PD	BL	ORB	PB	SEG	DS	PR	SP
LK																	
LSK	**0.73**																
WK	**0.52**	**0.42**															
WBK	**0.67**	**0.59**	**0.7**														
SC	**0.54**	**0.39**	**0.81**	**0.64**													
DEL	**0.65**	0.68	**0.65**	**0.68**	**0.57**												
WD	**0.65**	0.78	0.65	0.5	0.6	0.69											
SD	**0.47**	0.39	0.31	0.47	0.35	0.74	0.38										
ORD	0.17	0.22	0.18	0.25	0.11	0.64	0.06	0.39									
PD	0.61	0.57	0.72	0.74	0.59	0.83	0.54	0.52	0.33								
BLEND	**0.58**	0.8	0.47	0.62	0.51	0.73	0.78	0.48	0.31	0.55							
ORB	**0.51**	0.79	0.5	0.5	0.45	0.61	0.86	0.38	0.06	0.48	0.79						
PB	**0.53**	0.68	0.38	0.58	0.46	0.68	0.62	0.46	0.38	0.49	0.95	0.56					
SEG	0.28	0.51	0.22	0.1	0.21	0.36	0.47	0.13	0.1	0.31	0.42	0.34	0.4				
DS	**0.49**	0.67	0.42	0.37	0.41	0.63	0.7	0.45	0.16	0.52	0.62	0.66	0.56	0.66			
PR	**-0.62**	**-0.78**	**-0.55**	-0.47	-0.53	-0.69	-0.93	-0.45	-0.1	-0.54	-0.77	-0.83	-0.62	-0.56	-0.9		
SP	**0.6**	**0.42**	**0.82**	**0.74**	**0.72**	0.73	0.55	0.52	**0.25**	**0.8**	0.46	**0.47**	**0.39**	0.19	0.48	-0.54	
GARS	**-0.7**	**-0.8**	**-0.73**	**-0.8**	**-0.72**	**-0.82**	**-0.71**	**-0.53**	**-0.34**	**-0.78**	**-0.79**	**-0.66**	**-0.73**	**-0.51**	**-0.38**	0.73	-0.67

Note: LK=Letter Knowledge, LSK=Letter-Sound Knowledge, WK=Word Knowledge, WBK=Word Building Knowledge, SC=Sentence Completion, DEL=Deletion (Composite Score), WD=Word Deletion, SD=Syllable Deletion, ORD=Onset-Rime Deletion, PD=Phoneme Deletion, BLEND=Blending (Composite Score), ORB=Onset-Rime Blending, PB=Phoneme Blending, SEG=Phonemic Segmentation, DS=Dynamic Segmentation, PR=Prompts, SP=Spelling, AR-T=Global At-Risk Total Score. Bolded values are significant. The negative values for the Prompts scores indicate that as there are more prompts the performance is worse.

Exploratory factor analysis was performed including the subtests of the ODET. A principal component factor analysis with varimax rotation resulted in three-factors being retained (minimum eigen value of 1 for inclusion). The Bentler Comparative Fit Index was 1.00 indicating that the three-factor solution provided a robust and appropriate solution for the data. As can be seen in **Table 8**, the five subtests of Letter Knowledge, Word Knowledge Word Building Knowledge, Sentence Completion and Spelling all loaded

together with loadings of .87, .93, .81, .89 and .80, respectively. These subtests all involve some aspect of reading at a macro level rather than at the phonological level. It was anticipated that these tasks would be related as they involve reading of words and nonwords, spelling, and comprehension along with knowledge of letter names. The next

factor had loadings from Letter Sound Knowledge (.86), Onset-Rime Deletion (.50), Phoneme Deletion (.49), Phoneme Blending (.53), Phoneme Segmentation (.56), Dynamic Segmentation (.66), Prompts (-.47) and Word Building (.41). All of these tasks require phonological processing skills at a fairly micro level. The Word Building task is a task in which the child reads nonwords. As a result, the ability to decode and blend phonemes is required in addition to knowledge of letter sound relationships for successful performance on Word Building. The Dynamic Segmentation task has two values associated with it, the number correct and the number of prompts that were presented to the child. Better performance resulted in higher number correct and a lower number of prompts. As a result, Prompts was a variable that would be inversely related to number correct and therefore good performance on all of the other tasks. The third factor included Word Deletion (.77), Syllable Deletion (.40), Onset-Rime Blending (.83), Phoneme Blending (.58), Phoneme Segmentation (.48) and Prompts (-.51). These variables also represent phonological processing skills. The two variables that were loaded the most strongly on this factor, Word Deletion and Onset-Rime Blending ,represent larger units of phonological processing, deleting words from compound words and blending onsets and rimes. Blending tasks have been found to be tasks that younger children are able to perform well on before they are able to engage in deletion and segmenting tasks. The other variables that were loaded on this factor also represent phonological processing skills.

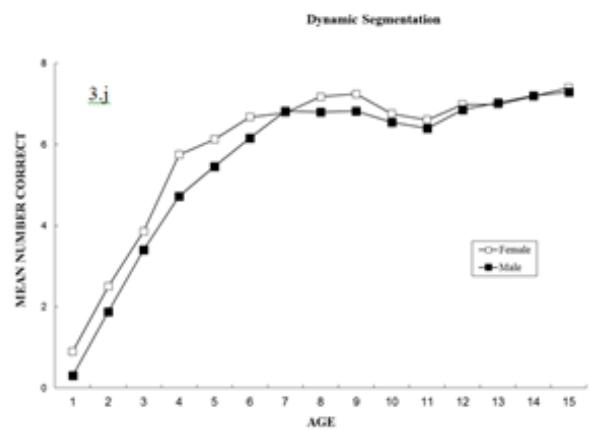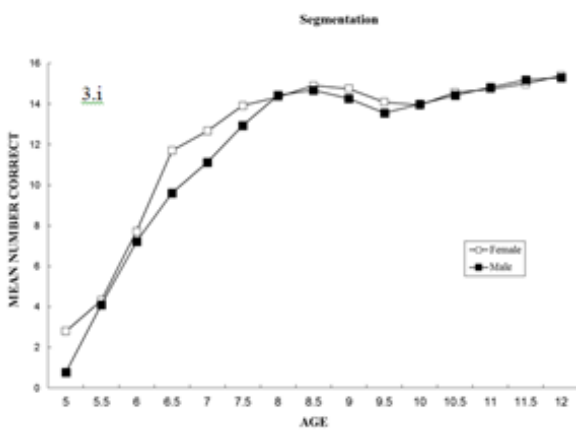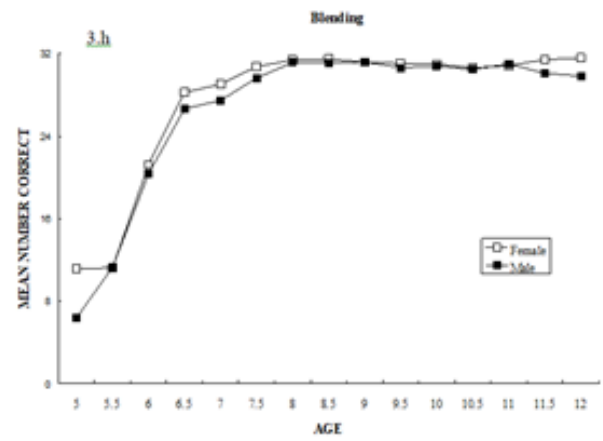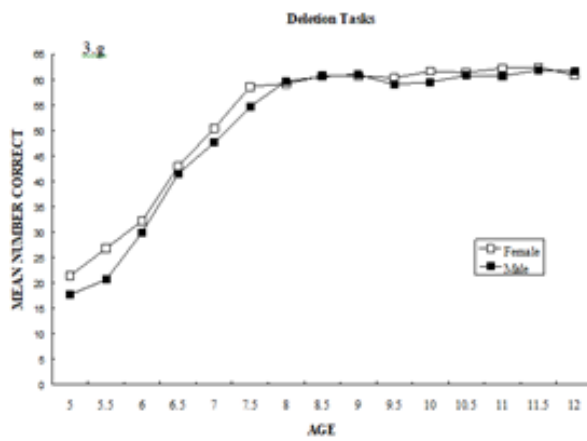**Table 8.** Factor Loading for the ODET by Factor.

ODET Subtask	Factor		
	1	2	3
Letter Knowledge	**0.87**	-0.02	0.08
Letter-Sound Knowledge	0.02	**0.86**	-0.09
Word Knowledge	**0.93**	0.05	0.15
Word Building Knowledge	**0.81**	0.41	0.16
Spelling	**0.80**	0.26	0.2
Sentence Completion	**0.89**	0.21	0.15
Word Deletion	0.18	-0.07	0.77
Syllable Deletion	**0.69**	0.27	**0.4**
Onset-Rime Deletion	**0.58**	**0.50**	0.27
Phoneme Deletion	**0.58**	**0.49**	0.28
Onset-Rime Blending	0.08	0.21	**0.83**
Phoneme Blending	0.22	**0.53**	**0.58**
Phonemic Segmentation	0.24	**0.56**	**0.48**
Dynamic Segmentation	0.39	**0.66**	0.35
Prompts (Dyn Seg)	**-0.46**	**-0.48**	**-0.51**

Interestingly but not surprisingly, the Letter Knowledge and Letter Sound Knowledge subtests loaded on different factors. Letter Knowledge was loaded with reading, spelling, and comprehension tasks and Letter Sound Knowledge was loaded with tasks that tapped phonological processing tasks. An individual can have a very good knowledge of letter names and be a nonreader, while a good reader must possess a working knowledge of the relationship between letters and the sounds that they represent. Letter name knowledge is orthogonal to reading while letter sound knowledge is critical.

The results of the exploratory factor analysis provided evidence regarding construct validity as the tasks that theoretically should have loaded together on the same factor did and those tasks that were theoretically different did not.

The standardization sample ranged in age from 5 to 12 years. The relationship between age and performance was curvilinear with the best performance at the older ages, but the progression depended on age. For example, on Letter Knowledge, the mean performance increased sharply from 5 years to 6.5 years and performance was nearly at maximum from 7 years of age through 12 years see **Figure 3a**. Letter Sound Knowledge had an even steeper increase from 5 to 6.4 years and was then asymptotic see **Figure 3b**. As can be seen in **Figures 3c-3f** the skills required for Word Knowledge, Word Building Knowledge, Sentence Completion and particularly Spelling took longer to develop. Similar patterns existed for the deletion tasks in which Word Deletion reached asymptote prior to Syllable Deletion, which reached asymptote prior to Onset-Rime Deletion, which reached asymptote prior to Phoneme Deletion. This pattern also existed for the blending and segmenting tasks see **Figures 3g-3k**. Given that the skills required to perform on these tasks take time to develop before reaching their potential, the differential

Letter Knowledge (3.a), Letter-Sound Knowledge (3.b), Sentence Completion (3.e), Spelling (3.f), Deletion Tasks (3.g), Blending (3.h), Segmentation (3.i), Dynamic Segmentation (3.j)
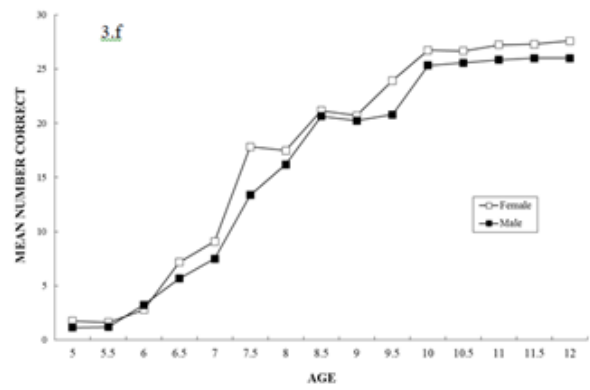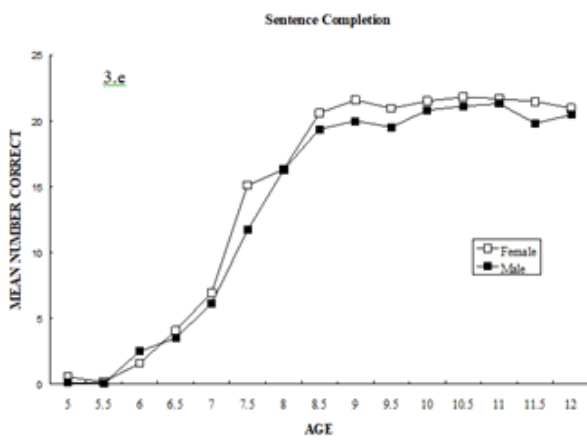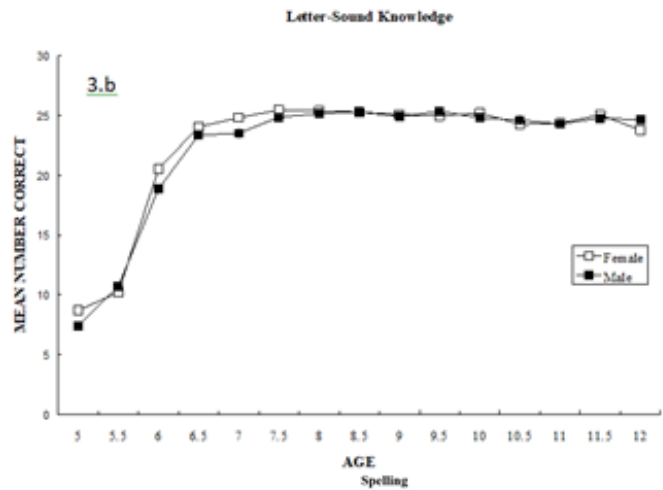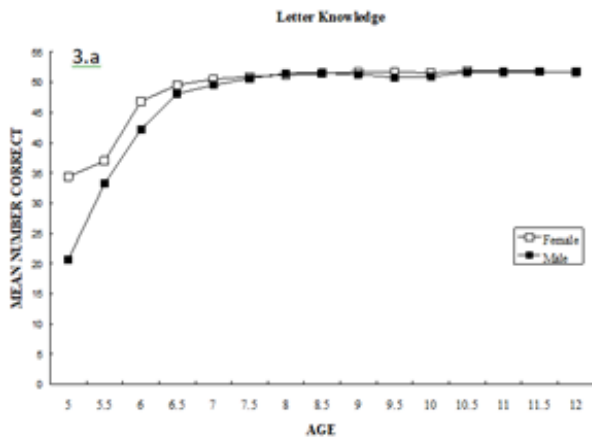
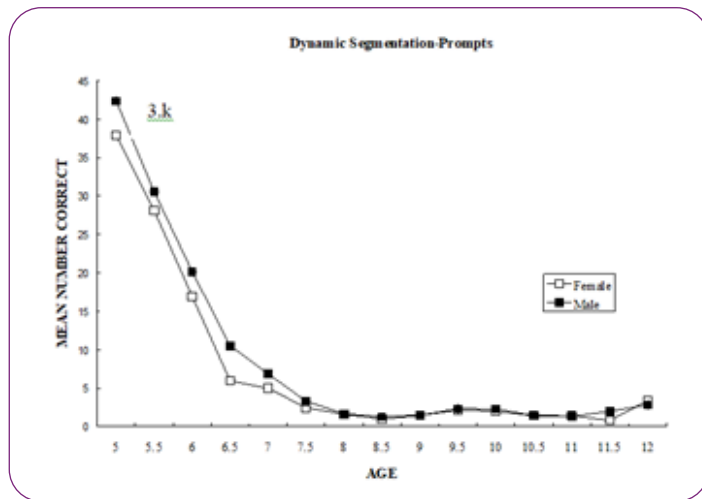performance on these tasks across age substantiates construct validity.

Figure 3 Means by Age, Gender and Subtask on ODET.

Construct validity was also evaluated by examining the performance of good and poor readers. Those participants who scored above a GARS of 14, which corresponded to the performing at or below the 25th percentile, were considered poor readers (M=25.18). Those who scored at or below a GARS of 14 were considered good readers (M=5.96). The participant's scores were entered into a 2 (At-Risk Group) x 11 (Subtask) repeated measures analysis of variance with repeated measures on subtask. The subtasks involved each of the subtasks of the ODET. The results indicated significant main effects of At-Risk group, $F(1, 36)$=22.08, p<.0001 and Subtask, $F(10, 360)$=470.03, p<.0001, and an interaction of At-Risk group and subtask, $F(10, 360)$=133.03, p<.0001. Post-hoc Student Newman Keuls analyses indicated that the At-Risk group significantly underperformed the Not At-Risk group on each subtest except for Letter Knowledge in which both groups performed at ceiling, which is not surprising as the participants in the validity study were third graders.

A similar analysis as described above was also carried out to determine if the poor reading group would perform significantly poorly on measures of phonological process (CTOPP-2), print knowledge (PPA), and reading (WIAT-III). The same good and poor readers were examined with a 2(A-Risk Group) x 11(Subtask) repeated measures analysis of variance with repeated measures on subtask. The results indicated significant main effects of At-Risk group, $F(1, 38)$=7.64, p<.008 and Subtask, $F(12, 456)$=55.86, p<.0001, and again an interaction of At-Risk group and subtask, $F(12, 456)$=1.82, p<.04. Post hoc Student Newman Keuls analyses indicated that the At-Risk group significantly underperformed the Not At-Risk group on each subtest with rather larger differences on the CTOPP-2's Elision and Segmenting Nonwords and the WIAT-III's Reading Comprehension, Word Reading and Word Reading Speed subtests. As was the case in the previous analysis, although the At-Risk group underperformed the Not At-Risk group, on the measures of the PPA, those differences did not reach significance for the Sound Symbol or Print Knowledge subtests (ps of .18 and .22, respectively). Students were performing nearly at ceiling

on the subtests evaluated print knowledge and sound symbol knowledge. Students who performed poorly on the ODET also scored poorly on the respective subtests of the CTOPP-2 and WIAT-III.

In each of the ways to evaluate validity, validity was strongly established. The ODET measures reading and its subcomponents, contains content that is consistent with the domain of interest, has subtasks that are correlated in the manner that suggests validity, is sensitive to age and to performance such that struggling readers can be identified.

Conclusion

The present study was undertaken to examine the psychometric properties of the ODET. The ODET has sufficient reliability as assessed by internal consistency, test-retest and inter-rater reliability. Validity was also found to be sufficient with examination of face, content, criterion, predictive and construct validities. The case for sufficient psychometric properties of the ODET was supported. Given the dearth of information to which parents have access with regard to reading difficulties and dyslexia, the ability to successfully administer an evaluation tool that provides them with useful and accurate information would prove to be quite helpful in their quest to find answers and to advocate for their children. The first step in the process would be to gain accurate information regarding their children's phonological processing, letter knowledge, reading, spelling and comprehension. The results of the present study indicate that the ODET has appropriate psychometric properties.

References

1. Norton ES, Beach SD, Gabrieli JD (2015) Neurobiology of dyslexia. Curr Opin Neurobiol 30(1):73-78.

2. Vellutino FR, Fletcher JM, Snowling MJ, Scanlon DM (2004) Specific reading disability (dyslexia): What have we learned in the past four decades? J Child Psychol Psychiatry 45(1): 2–40.

3. Lefly DL, Pennington BF (1991) Spelling errors and reading fluency in compensated adult dyslexics. Ann Dyslexia 41(1):141-162.

4. Lyon GR, Shaywitz SE, Shaywitz BA (2003) A definition of dyslexia. Ann. of Dyslexia 53(1): 1-14.

5. Cassar M, Treiman R, Moats L, Pollo TC, Kessler B (2005). How do the spellings of children with dyslexia compare with those of nondyslexic children? Read Writ 18(1): 27-49.

6. Sumner E, Connelly V, Barnett AL (2014) The influence of spelling ability on handwriting production: Children with and without dyslexia. J Exp Psycho: Learning, Memory, and Cognition 40(5): 1441–1447.

7. Carrion CA, Francks C, Fisher SE (2015) Evaluating the genetic risk for dyslexia in multi-generation families, Max Planck Institute Publications, Nijmegen, The Netherlands.

8. Eden G F, Olulade OA, Evans TM (2016) Developmental dyslexia. In: Gregory H, Steven LS (eds) Neurobiology of

Language. Academic Press, Washington DC, USA, pp. 815-826.

9. Shao S, Kong R, Zou L (2016) The roles of genes in the neuronal migration and neurite outgrowth network in developmental dyslexia: single-and multiple-risk genetic variants. Mol Neurobiol 53(6), 3967-3975.

10. Bruck M (1990) Word-recognition skills of adults with childhood diagnoses of dyslexia. Developmental psychology 26(3): 439-454.

11. Schatschneider C, Torgesen JK (2004) Using our current understanding of dyslexia to support early identification and intervention. J Child Neurol. 19(10): 759-765.

12. National Center for Educational Statistics (2019). The Nation's Report Card.

13. Tallal P, Miller SL, Bedi G (1996). Language Comprehension in Language-Learning Impaired Children Improved with Acoustically Modified Speech. Science 5: 81-84.

14. Krafnick AJ, Flowers DL, Napoliello EM (2011) Gray matter volume changes following reading intervention in dyslexic children. Neuroimage 57(3): 733-741.

15. Ozernov-Palchik O, Gaab N (2016) Tackling the 'dyslexia paradox': Reading brain and behavior for early markers of developmental dyslexia Wiley Interdiscip Rev Cogn Sci. 7(2): 156-176.

16. Gabrieli JD (2009) Dyslexia: A new synergy between education and cognitive neuroscience. Science 325(5938): 280-283.

17. Suggate SP (2016) A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. J Learn Disabil 49(1): 77-96.

18. Winkler CA (2016) Effective Interventions for Improving Reading Accuracy and Fluency in Children with Dyslexia. Scholarship and Engagement in Education 1(1): 2.

19. Bos C, Mather N (1999) Interactive, collaborative professional development in early literacy instruction: Supporting the balancing act. Learning Disabilities Research & Practice 14: 227-238.

20. Coyne MD, Kame'enui EJ, Simmons DC (2004) Improving beginning reading instruction and intervention for students with LD: Reconciling "All" with "Each." J Learn Disabil 37(3): 231-239.

21. Weiser B, Mathes P (2011) Using encoding instruction to improve the reading and spelling performances of elementary students At-Risk for literacy difficulties: A best-evidence synthesis. Review of Educational Research 81(2): 170-200.

22. Partanen M, Siegel LS (2014) Long-term outcome of the early identification and intervention of reading disabilities. Read Writ 27: 665–684.

23. McNulty MA (2003) Dyslexia and the life course J Learn Disabil 36(4):363-381.

24. Casey R, Levy SE, Brown K, Brooks-Gunn J (1992) Impaired emotional health in children with mild reading disability. J Dev Behav Pediatr 13(4):256-60.

25. Carroll JM, Iles JE (2006) An assessment of anxiety levels in dyslexic students in higher education. Br J Educ Psychol 76(3): 651-662.

26. Novita S(2016) Secondary symptoms of dyslexia: A comparison of self-esteem and anxiety profiles of children with and without dyslexia. Euro J Spl Needs Edu 31(2): 279-288.

27. Alexander-Passe N (2015) Investigating Post Traumatic Stress Disorder (PTSD) Triggered by the Experience of Dyslexia in Mainstream School Education. J Psychol Psychother 5(6): 1-6.

28. Alexander-Passe N (2015). Dyslexia: Investigating self-harm and suicidal thoughts/attempts as a coping strategy. J Psycho Psychother 5(6): 1.

29. McBride HEA, Siegel LS (1997) Learning disabilities and adolescent suicide. J Learn Disabil 30(6): 652-659.

30. Fuller-Thomson E, Hooper SR (2015)The association between childhood physical abuse and dyslexia: Findings from a population-based study. J Interpers Violence 30(9): 1583-1592.

31. Cortiella C, Horowitz SH (2014) The state of learning disabilities: Facts, trends and emerging issues. New York: National Center for Learning Disabilities.

32. Gonzalez JM, Cannell MB (2015) Disproportionate prevalence rate of prisoners with disabilities. J Disability Policy Stu 27(2): 106-115.

33. Mears DP, Aron LY (2003). Addressing the needs of youth with disabilities in the juvenile justice system: The current state of knowledge, The Urban Institute, Washington DC, USA.

34. Guttorm TK, Leppänen PHT, Tolvanen A, Lyytinen H (2003) Event-related potentials in newborns with and without familial risk for dyslexia: principal component analysis reveals differences between the groups. J Neural Transm (Vienna) 110(9): 1059-1074.

35. Leppänen PHT, Hämäläinen JA, Guttorm TK (2012). Infant brain responses associated with reading-Related skills before school and at school age. Neurophysiol Clin 42(1-2): 35-41.

36. Scarborough HS (1998) Early identification of children At-Risk for reading disabilities: Phonological awareness and some other promising predictors. In: Shapiro BK, Accardo PJ, Capute AJ, editors. Specific reading disability: A view of the spectrum. Timonium, MD: York Press. 75–119.

37. Lonigan CJ, Burgess SR, Anthony JL (2000) Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. Dev Psychol 36(5): 596-613.

38. Wagner, Torgesen, Rashotte, Pearson (2014) The comprehensive test of phonological processing-second Edition (CTOPP-2;) Canadian J School Psycho 30(2): 155-162.

39. Tennant KE (2014) Test Review: Comprehensive test of phonological processing–second Edition (CTOPP-2). J Psyschoedu Assess, 32(7): 678-681.

40. D. Wechsler (2009) Wechsler Individual Achievement Test (3rd edn), NCS Pearson, San Antonio, USA.

41. McCrimmon AW, Climie EA (2011) Test review: Wechsler individual achievement test-third edition. Canadian J Sch Psycho 26(2): 148-156.

42. Williams K (2014) Phonological and Print Awareness Scale, Western Psychological Services, Canada.

43. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J Royal Stat Society, Series B (Methodological)39(1): 1-38.

44. Wagner RK, Torgesen JK (1987) The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological bulletin 101(2): 192-212.

45. Snowling MJ (1995) Phonological processing and developmental dyslexia. Journal of research in reading 18(2): 132-138.

46. Vygotsky LS (1978) Mind in society: The development of higher Psychological Processes, Harvard University Press, London, England.

47. Wood D, Bruner JS, Ross G (1976). The role of tutoring in problem solving. J child psycho and psychiat 17(2), 89-100.