Assessing written language in Greek: Validity and reliability considerations

Ioanna Talli

Aristotle University of Thessaloniki, Greece

*Corresponding author: Ioanna Talli, Aristotle University of Thessaloniki, Mantzaridi 3, Panorama, Thessaloniki, 55236 Greece.

Email: tallijoan@yahoo.com

Phone number: +30 2310 346845

Abstract

This study aims at contributing to the reading assessment tools for Greek speaking population by evaluating the validity and reliability of decoding and reading comprehension measures. We tested 173 monolingual Greek children (aged 7-10) from elementary school Grades 2, 3, and 4, which were divided into three age groups: (i) 7-8 years, (ii) 8-9 years, (iii) 9-10 years. We employed the following measures: Word decoding measures (WDM): (i) a word and a non-word reading test and (ii) a reading fluency test. Reading comprehension measures (RCM): (i) a standardized reading comprehension test and (ii) a reading comprehension test based on the French test ECoSSe (*Epreuve de Compréhension Syntaxico-Sémantique* [Test of syntactico-semantic comprehension]).We investigated the convergent validity of these measures and found significant and strong (S & S) correlations between all these measures for the 7-year-old children; for the 8-year-old and 9-year-old groups S & S correlations were only found within each measure category. Furthermore, to investigate the external consistency of these measures we performed test-retest reliability and administered them for a second time after two weeks to a smaller group of 30 individuals (10 for each grade) of this cohort. Correlation analysis indicated S & S correlation. We also investigated internal consistency by performing split-half reliability, which indicated very high estimate for all tests except the second RCM. These results show the interdependence of different aspects of reading abilities in the early stage of reading development.

Keywords

reading assessment, decoding, reading comprehension, reliability, validity

1. Introduction

Reading acquisition is a challenging process for the early school age children as it compromises different kind of skills, in particular, decoding and reading comprehension skills (Perfetti & Hogaboam, 1975). While word decoding and written word identification requires skills specific to reading, reading comprehension requires both the ability to accurately and fluently identify written words and the ability to comprehend language in general (Florit & Cain, 2011; Gough & Tunmer, 1986; Stanovich, 1986). More specifically, word decoding requires phonological abilities such as grapheme to phoneme conversion (Castles & Coltheart, 2004; Ehri, 1998). On the other hand, reading comprehension, in accordance with current views in the field (Cain, Oakhill & Bryant, 2004; Kendeou, van den Broek, White & Lynch, 2009), requires two level cognitive processes: (i) *lower level processes* including translation of the written input into language units that have meaning, i.e. decoding (Stanovich, 2009), reading fluency (Kuhn, & Stahl, 2003), and vocabulary knowledge (Nagy, Anderson & Herman, 1987) and, (ii) *higher level processes* consisting of transforming these units into a mental representation that has meaning. However, reading comprehension is affected by a wide range of cognitive skills and processes (Kendeou, Broek, Helder, & Karlsson, 2014; Cain & Oakhill, 2011; Vellutino, Tunmer, Jaccard & Chen, 2007) and not all reading comprehension tests are tapping into the same cognitive processes (Kendeou, Papadopoulos & Spanoudis, 2012).

To assess the children's reading abilities and provide the proper intervention for poor readers, specialised tools of reading assessment have been developed (for English: WORD; Rust et al., 1992; Test of Word Reading Efficiency [TOWRE]; Torgesen, Wagner, & Rashotte, 1997; Neale Analysis of Reading Abilities (NARAII; Neale, 1989; for French 'Alouette',

Lefavrais, 1967; ODEDYS 2 - Outil de DEpistage des DYSlexies, Jacquier-Roux, Valdois, Zorman, 2005; Test de Lecture pour Cours Préparatoire [TLCP], Pasquier, 1979; Epreuve d'évaluation de la compétence en lecture [LMC-R], Khomsi, 1999).

Recently, the development of reading abilities of Greek children has intensively attracted research attention (e.g. Anastasiou & Protopapas, 2015; Desrochers, Manolitsis, Gaudreau, & Georgiou, 2017; Diamanti, Mouzaki, Ralli, Antoniou, Papaioannou & Protopapas, 2017; Nikolopoulos, Goulandris, Hulme & Snowling, 2006; Porpodas, 2012; Protopapas, Simos, Sideridis, & Mouzaki, 2012). Researchers have focused on how Greek children acquire reading and investigated the language specific effects on reading acquisition in order to develop assessment tools. Greek poses specific challenges to the reading acquisition research as it has a relatively transparent orthography. Specifically, it occupies the second position according to Seymour's et al. (2003) classification, after Finnish. It is, thus, of great importance to examine how the reading acquisition process is realized in Greek and whether the acquisition of such a transparent language still poses learning challenges for children with typical development as well as for children with reading disabilities (Georgiou, Parrila, & Papadopoulos, 2008; Georgiou, Torppa, Manolitsis, & Lyytinen, 2012; Protopapas & Skaloumbakas, 2007; for a review see Protopapas, 2017). Accurate description of the difficulties that Greek children face with learning to read can be provided by specialized measures of the different cognitive processes involved in reading development.

While little work has been performed up to now on assessing reading abilities in Greek, the available materials provide measurement of different aspects of the reading ability (decoding and/or comprehension or fluency). We provide a brief summary of them below:

Reading comprehension measures

(i) Screening Test of Reading Ability (STRA; Τεστ Ανίχνευσης της Αναγνωστικής Ικανότητας) (Tafa, 1995). This test mostly assesses the child reading comprehension abilities. It has been developed to be employed for children aged 6;9-10;1 years. In particular, it is a standardized multiple-choice sentence completion test consisted of 42 sentences.

(ii) Lamda test for the detection of learning skills and difficulties, Grades 2nd-4th and 5th-8th (ΛΑΜΔΑ, Λογισμικό Ανίχνευσης Μαθησιακών Δεξιοτήτων και Αδυναμιών; Skaloumbakas & Protopapas, 2007). This test has been developed to assess eight domains, among which are reading comprehension, morpho-syntactic processing and working memory. The other domains are spelling, oral language comprehension, vocabulary, non-verbal IQ and perception of music characteristics.

Reading comprehension and decoding

(iii) Screening and investigation test of reading disabilities in kindergarten and 1st and 2nd Grade (Εργαλείο Ανίχνευσης και Διερεύνησης των Αναγνωστικών Δυσκολιών στο Νηπιαγωγείο και Α - Β´ Δημοτικού; Porpodas, 2007). This test has been developed to assess the reading abilities of children aged 5-7 years and identify the reading difficulties they may have. More specifically, it evaluates the following domains: (i) phonological awareness, (ii) short-term memory, (iii) decoding and (iv) reading comprehension by providing nine subtests for all of them in total.

Multiple assessment of reading abilities

(iv) Reading test (Test-A) (Τεστ ανάγνωσης: ΤΕΣΤ-Α) (Panteliadou & Antoniou, 2007): This test has been developed to assess the reading abilities of children aged 8-15 and identify the reading difficulties they may have. More specifically, it evaluates the following domains: (1) word decoding, (2) fluency, (3) morphology and syntax, and (4) reading comprehension, by providing ten subtests for all of them in total.

(v) Reading test ΔΑΔΑ (Διερεύνηση Αναγνωστικών Δεξιοτήτων και Αδυναμιών) (Panteliadou, Antoniou, & Sideridis, 2018): This test assesses the reading abilities of elementary school-aged children and it evaluates: (i) decoding (word and pseudoword reading, word and pseudoword discrimination and word identification), (ii) reading fluency and (iii) reading comprehension.

2. The present study

This study aims at contributing to the field of reading assessment in Greek by examining the validity and reliability of new measures assessing decoding and reading comprehension. The development of these materials was highly motivated by the need to assess children with language and reading disorders. These materials proved to successfully distinguish children with reading and language disorders from children with typical development. All measures but one reading comprehension measure (i.e. STRA standardized test) were developed or adjusted to Greek by Talli (2010). The purposes of the present study is to further examine their appropriateness for testing written language abilities in Greek speaking children from Grade 2 to Grade 4 (ages 7-10), by examining their validity and reliability, which constitute crucial qualities of psychometric measures.

A valid test measures well what it has been designed to measure. A reliable test generates stable and consistent results (Putka & Sackett, 2010). Once the validity and the reliability criteria are met, then actual variations in decoding and reading comprehension can be considered as reflecting differences in the reading ability and not accidental variations within the tested sample.

The specific questions the present study addresses are the following:

1. How are word reading and reading comprehension measures associated among them? Is there a relationship both within and between each category measure?
2. Do WRM and RCM increase as a function of age level?
3. Can scores on the different reading measures predict age level?
4. Are WRM and RCM stable across the two categories?

3. Method

3.1 Participants

One hundred and seventy-three students from Grade 2 to Grade 4 were recruited from public primary schools in city of Thessaloniki, Greece. Only those children whose parents gave written permission to participate in the research were included in the study. None of them had a history of speech and language problems, or a referral to a speech therapist or a history of reading and writing problems. Only children who demonstrated normal progress in school were included. Furthermore, any child who scored below the 25th percentile in the non-verbal IQ test (Raven's Coloured Progressive Matrices) was excluded from the study.

3.2 Procedure

Participants were assessed at their school individually in one or two sessions of a total duration of 1 hour and 10 minutes.

3.3 Materials

Word decoding measures (WDM)

Word and non-word reading. The number of items correctly read in one minute was measured in two tasks. One task included 50 high-frequency regular words selected on the basis of lexical frequency (Hellenic National Corpus,HNC [http://hnc.ilsp. gr]). The other included 50 pseudowords matched with the words for the: number of phonemes, number of syllables and mean syllable frequency (both for stressed and unstressed syllables) within each word ("ILSP PsychoLinguistic Resource" [IPLR]; Protopapas et al., 2012). Errors of stress were taken into account when calculating words and pseudowords correctly read.

Reading fluency ('Giro Giro oli', adaptation of 'Alouette', Lefavrais, 1967; Talli, 2010; Talli, Stavrakaki & Sprenger-Charolles, 2015). The children were each asked to read aloud a 271-word text as accurately and rapidly as possible. The text included rare words and numerous pieces of misleading contextual information (e.g., 'korakia' [crows] instead of 'koralia' [corals], after 'limni' [lake] and 'neron' [waters]). For more details on the characteristics of this test see Talli, Stavrakaki and Sprenger-Charolles (2015). This test therefore assessed decoding skills rather than word-in-context reading skills. The reading score was calculated by adding the total number of non-corrected errors and the total number of non-read words to the total reading time (calculated with a stopwatch in seconds, with a limit of 180 seconds).

Reading comprehension measures (RCM)

Screening Test of Reading Ability (STRA). This was a standardized multiple-choice sentence completion test (Screening Test of Reading Ability; Tafa, 1995) consisted of 42 sentences. The child was required to read silently the sentences and choose the word that was missing in each sentence, between four words that were given below each sentence. This test assesses mostly the lexical-semantic skills and this is why we chose to include it to our battery, since our second reading comprehension task assessed syntactico-semantic skills. The duration of this test was 40 minutes. Raw scores were taken into account.

Reading comprehension. This test was based on the French test ECoSSe (*Epreuve de Compréhension Syntaxico-Sémantique* [Test of syntactico-semantic comprehension], Lecocq, 1998; Talli, 2010 for the Greek version). More precisely, the Greek version contained seven types of utterances (5 in each category). These types of utterances are as following: (1) Utterances with a negation and a connective (e.g., *Το λουλούδι είναι κίτρινο αλλά όχι το πουλί* [the flower is yellow but not the bird]); (2) Utterances with "neither...nor" (e.g., *Το αγόρι δεν είναι ούτε ψηλό ούτε ξανθό* [the boy is neither tall nor blond]); (3) Utterances with spatial terms (e.g., *Η μωβ κούπα είναι κάτω από το μπολ* [the purple cup is under the bowl]); (4) Reversible passives (5-6) Two types of utterances with a relative clause (subject and object), connected to the first noun (*Ο κυνηγός που κουβαλάει τον στρατιώτη έχει ένα μάτι* [the hunter who carries the soldier has one eye] versus *Ο ψαράς που σπρώχνει ο δύτης είναι λυπημένος* [the fisherman whom the diver is pushing is sad]); and (7) Utterances with accusative direct object clitics (e.g., *Ο λαγός τον χτυπά* [the rabbit is hitting him]). Every child was asked to read each utterance, written on one page. Then s/he was shown a page with four pictured choices, and asked to select the correct one (without going back to the page on which the utterance was written). The percentage of errors was recorded and analyzed.

4. Results

Means and standard deviations for word decoding measures (word reading, non-word reading and reading fluency), as well as reading comprehension measures (STRA and Greek Ecosse) are presented as a function of age level in **Table1**.

Table 1. Decoding and reading comprehension skills: Means (and SDs) of the three groups

Age group	Group 1 (7-8 yrs)	Group 2 (8-9 yrs)	Group 3 (9-10 yrs)
	7;6 (4,58)	8;7 (3,58)	9;6 (3,12)
Word identification tests			
i Word reading per minute	57,78 (20,37)	75,47 (16,56)	80,42 (20,32)
ii Non-word reading per minute	36,59 (11,70)	46,00 (12,60)	51,40 (13,66)
iii Reading fluency (composite score)	325,08 (41,47)	283,26 (46,69)	248,65 (54,41)
Reading comprehension tests			
a. STRA raw score	24,33 (6,76)	30,47 (6,65)	35,06 (3,97)
b. ECOSSE Greek version (percentage of errors)	24,12 (9,85)	20,99 (8,20)	18,62 (8,70)

4.1 Validity analysis

We performed validity analysis to investigate how word reading and reading comprehension measures are associated among them and whether there is a relationship both within and between each category measure. Specifically, we investigated the convergent validity of the above measures by performing *Pearson* correlation coefficient analysis for each of the age groups. **Table 2** shows the correlations between measures for each of the three groups. For the 7-year-old children all the correlations were significant. We found very strong correlations within all WDM (r = .77 between word and non-word reading, r = -.82 and -.84 for word and non-word reading respectively in relationship with reading fluency; $p < .001$) and strong correlation within RCM (r = -.53, $p <.001$).As regards the relationship between WDM and RCM, the strongest correlations were found between STRA and reading fluency measures (r = -.061, $p < .001$) and between STRA and word and non-word reading measures (r = .54 and .51 respectively, $p < .001$). The Greek *Ecosse* correlated strongly with reading fluency (r = .45, $p < .001$), but moderately with word and non-word reading (r = -.32 and -.38; $p < .012$ and $p < .003$ respectively).

Table 2. Correlations between measured variables by age group

TEST CATEGORIES	*Word identification*			*Reading comprehension*	
TEST SUBCATEGORIES	**WR**	**NWR**	**RF**	**STRA**	**ECOSSE[GR]**
7-8 year-old					
Word identification tests					
Word reading (WR)	1	,773**	-,821**	,538**	-,321*
Non-word reading (NWR)	,773**	1	-,840**	,514**	-,380**
Reading fluency (RF)	-,821**	-,840**	1	-,609**	,448**
Reading comprehension tests					
STRA	,538**	,514**	-,609**	1	-,530**
ECOSSEGR	-,321*	-,380**	,448**	-,530**	1
8-9 year-old					
Word identification tests					
Word reading	1	,659**	-,634**	,316*	-,023
Non-word reading	,659**	1	-,727**	,379**	-,152
Reading fluency	-,634**	-,727**	1	-,389**	,067
Reading comprehension tests					
STRA	,316*	,379**	-,389**	1	-,490**
ECOSSEGR	-,023	-,152	,067	-,490**	1
9-10 year-old					
Word identification tests					
Word reading	1	,664**	-,386**	,145	,062
Non-word reading	,664**	1	-,637**	,108	,036
Reading fluency	-,386**	-,637**	1	,108	,036
Reading comprehension tests					
STRA	,145	,108	-,364**	1	-,466**
ECOSSE[GR] version	,062	,036	,240	-,466**	1

**Correlation is significant at the 0.01 level (2-tailed) *Correlation is significant at the 0.05 level (2-tailed)

For the 8-year-old group, significant and strong correlations were only found within each measure category (only within WDM or RCM), while only one of the two RCM correlated with WDM. More specifically, word and non-word reading correlated strongly (r =.66, $p < .001$), so did reading fluency with both the other decoding measures (r = -.63 and -.73 for word and non-word reading respectively; $p < .001$). Strong correlation was also found within RCM (r = -.49, $p < .001$). As regards the relationship between WDM and RCM, only STRA correlated moderately andsignificantly with all three WDM (word and non-word reading: r = .32 and .38 respectively, $p < .003$; reading fluency: r = .39, $p < .003$). None of the correlations between *Ecosse* and the three WDM was significant.

Finally, for the oldest group significant and strong correlations were only found within each measure category, while there was only one significant correlation between WDM and RCM. More specifically, word and non-word reading correlated strongly (r =.66, $p < .001$), so did reading fluency with non-word reading (r = -.64, $p < .001$), while it correlated moderately with word reading (r = -.39; $p = .004$). The two RCM correlated significantly and strongly (r = -.47, $p < .001$). As regards the relationship between WDM and RCM, only STRA correlated moderately andsignificantly with one WDM, the reading fluency measure (r = -.37, $p < .007$). None of the other correlations were significant.

To further investigate the validity of WRM and RCM we explored whether the relevant scores increase as a function of age level. In particular, five repeated measures analyses of variance (ANOVAs) were conducted on word reading, non-word reading, reading fluency, STRA and Greek Ecosse (within-subjects factor) with Group (1, 2 and 3) as the between-subjects factor. These analyses were followed by post hoc comparisons (Tukey's HSD test, $p < .05$).

WDM. For word and non-word reading the group effect was significant ($F_{(2, 170)}$ = 22.62 and 20.41 respectively, $p < .001$). Post-hoc comparisons indicated that for both word and non-word reading group 1 had significantly worse performance than group 2 and group 3, but group 2 did not have significantly worse performance than group 3. For reading fluency the group effect was significant $F_{(2, 170)}$ = 37.31, $p < .001$. Post-hoc pairwise comparisons using Tukey's HSD test showed that there was a significant difference in the performance of the three groups (group 1 < group 2 < group 3).

RCM. For STRA the group effect was significant $F_{(2, 170)}$ = 46.58, $p < .001$. Post-hoc Tukey comparisons revealed that there was a significant difference in the performance of the three groups (group 1 < group 2 < group 3). For the Greek adaptation of Ecosse the group effect was significant ($F_{(2, 170)}$ = 5.46, $p = .005$). Post-hoc comparisons indicated that significant difference existed only between group 1 and group 3 (group 1 < group 3). There was no significant difference between group 1 and group 2 or between group 2 and group 3.

Table 3. Results of linear regression analyses: effect of the performance at different reading measures on age group.

Variable		B	Std. Error	t	p	R^2
1	(Constant)	,016	,208	,079	,937	,352
	STRA	,065	,007	9,628	,000	
2	(Constant)	1,855	,508	3,650	,000	,406
	STRA	,045	,008	5,389	,000	
	Reading fluency	-,004	,001	-3,934	,000	

Note: Dependent Variable: age group. Word and nonword reading were calculated individually as correct items per minute (words, nonwords). Reading fluency was calculated as a composite score of total errors, total words non read and total time in sec. (with a limit of 180 sec.)

4.2 Regression analysis

We further provided validation for these reading measures by examining whether they can predict the participant chronological age. We thus performed a linear regression analysis with the stepwise method, with scores on the five reading measures (reading fluency test, word reading, non-word reading, STRA, Ecosse Greek version) as an independent variable accounting for differences in age (in months), which was the dependent variable (**Table 3**). Results showed that performance on STRA and reading fluency test contributed significantly to age: $F_{(1, 171)}$ = 92.71, $p = .000$, for STRA as a predictor, with an R^2 of .352 and $F_{(2, 170)}$ = 58.02, $p = .000$, for both STRA and reading fluency test as predictors, with an R^2 of .406. It is estimated that 35% of the age groups can be accounted for by performance on STRA, while 41% of the age groups can be accounted for by performance on both STRA and reading fluency test.

4.3 Reliability analysis

To investigate whether WRM and RCM are stable across the two categories we performed reliability analysis and we explored both internal and external consistency. In particular, to explore the internal consistency of the employed measures (WDM and RCM) we performed split-half reliability analysis (Spearman-Brown and Guttman Split-Half coefficient) and administered them to a smaller group of individuals (10 for each grade: 30 children in total). We did not perform reliability test for the

STRA test, since it is standardized and it has been already performed such analysis (see Table 4). We then divided the items of each test into two subtests (odd-even method), each half the length of the original test and we computed Spearman-Brown correlation of the two halves of the test. The Spearman-Brown formula provides an estimate, based on the split-half correlation, of the reliability of the test as a whole. In Table 4, Spearman-Brown as well as Guttman coefficients are reported. The reading fluency test showed the greatest reliability, the estimate being .99, followed by non-word reading (.94) and word reading (.93). The Ecosse Greek adaptation test had the lowest reliability (.41). This was presumably because this test assesses seven different types of utterances with a certain degree of syntactic complexity (for example passives), which were perhaps difficult to comprehend, at least for the age range we tested. If we had administered it to older children, we may have yielded most consistent results.

To investigate the external consistency of the above measures and to explore whether they are stable over time and yield the same results, we performed test-retest reliability analysis. We administered the reading measures for a second time after a period of two weeks to a smaller group of children (30 children, 10 for each age group). For word reading, reading fluency and STRA score correlation coefficient for test–retest reliability was .99, while for non-word reading .98 and for Greek adaptation of Ecosse was .96. Consequently, all reading measures employed in this study show high test-retest reliability (see **Table 4**).

Table 4. Split-half reliability (Spearman-Brown and Guttman coefficients) and test-retest reliability of all reading measures.

Measure	Half A M (SD)	Half B M (SD)	N	Spearman-Brown coefficient	Guttman Split-Half coefficient	Test-retest reliability correlation
Word reading	23,87 (1,85)	23,66 (2,11)	50	,93	,93	,98
Non-word reading	21,36 (3,78)	20,27 (4,09)	50	,94	,94	,98
Reading fluency	82,47 (25,33)	80,80 (25,91)	271	,99	,99	,99
STRA	-	-	42	-	,93	,99
ECOSSE[GR] version	14,53 (1,70)	14,27 (1,26)	35	,41	,39	,96

5. Discussion

The current study aimed at evaluating the validity and reliability of reading measures tapping upon different aspects of reading, in particular word decoding and reading comprehension in a consistent orthography, namely Greek.The specific questions that were set were: First, if there is a relationship both within and between each category measure, second, if WRM and RCM increase as a function of age level, third if scores on reading measures can predict age level and fourth if WRM and RCM are stable across the two categories. The results showed that for the 7-year-old children there were significant and strong (S & S) correlations within the WDM and RCM categories. We found very strong correlations within all WDM and strong correlation within RCM. Regarding the relationship between WDM and RCM, the strongest correlations were found between STRA and reading fluency measures and between STRA and word and non-word reading measures. The Greek adaptation of Ecosse correlated strongly with reading fluency, but moderately with word and non-word reading. On the other hand, for the 8-year-old group S & S correlations were only found within each measure category (only within WDM or RCM), while only one of the two RCM correlated with WDM (only STRA correlated moderately and significantly with all three WDM). Finally, again for the 9-year-old group S & S correlations were only found within each measure category and not between WDM and RCM, since only STRA correlated moderately andsignificantly with the reading fluency measure.

Concerning the second question, whether the relevant scores increase as a function of age level, we obtained significant differences in the performance of the three groups for the reading fluency test and STRA test. For word and non-word reading tests significant differences were only found between group 1 and groups 2 and 3 and for the Greek adaptation of Ecosse only between groups 1 and 3. Consequently, it can be concluded that scores in reading fluency test increase as a function of age level. However, for word and non-word reading tests scores increase as a function of age level but only for the first two groups. The older group did not perform much better than the second. This may be because the performance in decoding is not very different among older children who have already acquired such reading skills (8-9 and 9-10 years), while the younger group (7-8 years) have acquired as well decoding skills but are not such good decoders yet (Byrne, 2014; Stanovich, Cunningham & Feeman, 1984). As far as the reading comprehension test (Greek adaptation of Ecosse) is concerned, performance was ameliorated as a function of age level, but this better performance did not reach the level of significance. Significance may be obtained later on in older age.

Concerning the third question, the results of the regression analyses showed that reading performance can discriminate children

as it can predict chronological age. Consistent with previous reports in the literature (e.g. Protopapas, Mouzaki, & Sideridis, 2007), our data showed that reading fluency test and a reading comprehension test (STRA) accounted for a substantial proportion of age group. Apparently, reading comprehension (STRA) and reading fluency are significant predictors of age and are the measures that are most substantially and consistently improving by age. The other reading comprehension test (Greek adaptation of Ecosse) did not demonstrate age predictability presumably due to the specific syntactic demands of this test (e.g. inclusion of passive sentence whose acquisition develops slowly). It is a matter of further research to investigate whether this test's predictability arises at a later stage.

To address the fourth question, we explored both internal and external consistency by conducting split-half reliability on the one hand and test-retest reliability on the other hand. The results of the split-half reliability analysis yielded high reliability (over .93) for all but one test (the Greek adaptation of Ecosse) and the results of the test-retest reliability to a subgroup of 30 children yielded high reliability for all measures for all three groups (over .88). Consequently all decoding measures are reliable, at least concerning their external consistency and concerning their internal consistency, all but one test are reliable (the Greek adaptation of Ecosse). This test assesses seven different types of utterances and for the age range that we tested, children did not demonstrate consistent results probably because it was difficult for them. Perhaps this test would have higher internal consistency for older children.

In addition the validity analysis indicated S & S correlations even between measures tapping upon different reading abilities, namely word decoding measures and reading comprehension measures for the younger group while such correlations were only established within the measure category for the older group of the children. These results show the interdependence of different aspects of reading abilities in the early stage of reading development. Collectively, the results from the present study indicate that both WDM (word and non-word reading and reading fluency) and RCM (at least one of them) employed in this study are reliable and valid measures of reading abilities in Greek.

Acknowledgements

References

Anastasiou, D., & Protopapas, A. (2015). Difficulties in lexical stress versus difficulties in segmental phonology among adolescents with dyslexia. *Scientific Studies of Reading*, 19(1), 31-50.

Byrne, B. (2014). *The foundation of literacy: The child's acquisition of the alphabetic principle*. Psychology Press.

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of learning disabilities*, 44(5), 431-443.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of educational psychology*, 96(1), 31-42.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read?. *Cognition*, 91(1), 77-111.

Desrochers, A., Manolitsis, G., Gaudreau, P., & Georgiou, G. (2017). Early contribution of morphological awareness to literacy skills across languages varying in orthographic consistency. *Reading and Writing*, 1-25.

Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Frontiers in psychology*, 8, 2039.

Ehri, L. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies?. *Educational Psychology Review*, 23(4), 553-576.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, 100(3), 566-580.

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing*, 25(2), 321-346.

Gough, P.B., & Tumner, W.E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7, 6–10.

Hulme, C., & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal

study. *Journal of Experimental Child Psychology*, 94(1), 1-17.

Jacquier-Roux, M., Valdois, S., & Zorman, M. (2005). Outil de dépistage des dyslexies (ODEDYS). *Grenoble: Laboratoire des sciences de l'éducation*.

Kendeou, P., Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10-16.

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction*, 22(5), 354-367.

Kendeou, P., Van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of educational psychology*, 101(4), 765-778.

Khomsi, A. (1999). *Epreuve d'évaluation de la compétence en lecture: lecture de mots et compréhension-révisée*. Les Editions du Centre de psychologie appliquée.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of educational psychology*, 95(1), 3-21.

Lefavrais, P. (1967). *Test de l'Alouette: test d'analyse de la lecture et de la dyslexie*. Éditions du Centre de psychologie appliquée.

Lesch, M. E, & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 285-294.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal*, 24(2), 237-270.

Neale, M.D. (1989). Neale analysis of reading ability: Revised British edition. Windsor, UK: NFER.

Nikolopoulos, D., Goulandris, N., Hulme, C., & Snowling, M. J. (2006). The cognitive bases of learning to read and spell in Greek: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, 94(1), 1-17.

Panteliadou, S., & Antoniou, F. (2007). Test Alpha (Reading test). *ΕΠΕΑΕΚ, ΥΠΕΠΘ.(Standardised test for reading disabilities in Greek)*.

Panteliadou, S., Antoniou, F., and Sideridis, G. (2018). *Διερεύνηση αναγνωστικών δεξιοτήτων και αδυναμιών [Investigation of Reading Skills and Weaknesses]*. Nicosia: Rocket-lexia.

Pasquier, D. (1979). Test de lecture pour le Cours Préparatoire-TLCP. *Issy-les-Moulineaux: Editions Scient (ques et Psychotechniques)*.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology*, 67(4), 461-469.

Porpodas, C. D. (2012). The way Greek children are taught to read and spell in primary. *Reading and Writing Disorders in Different Orthographic Systems*, 52, 177. Chicago.

Protopapas, A. (2017). Learning to read Greek. In L. T. W. Verhoeven & C. A. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 155–180). Cambridge University Press.

Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology*, 33(3), 217-240.

Protopapas, A., & Skaloumbakas, C. (2007). Traditional and computer-based screening and diagnosis of reading disabilities in Greek. *Journal of Learning Disabilities*, 40(1), 15-36.

Putka, D. J., & Sackett, P. R. (2010). Reliability and validity. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 9-49). New York, NY, US: Routledge/Taylor & Francis Group.

Rust, J., Golombok, S., & Trickey, G. (1992). Wechsler Objective Reading Dimensions. London: The Psychological Corporation.

Seymour, P. H., Aro, M., Erskine, J. M., & collaboration with COST Action A8 network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, 94(2), 143-174.

Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of education*, 189(1-2), 23-55.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading research quarterly*, 360-407.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Relation between early reading acquisition and word decoding with and without context: A longitudinal study of first-grade children. *Journal of Educational Psychology*, 76(4), 668-677.

Talli, I., Stavrakaki, S., Sprenger-Charolles, L. (2015). Assessment of reading abilities in Greek: The Greek version of the

French reading test Alouette. *Studies for Greek Language (Μελέτες για την Ελληνική Γλώσσα)*, 35, 523-532. http://ins.web.auth.gr/images/MEG_PLIRI/MEG_35_523_532.pdf

Torgesen, J., Wagner, R., & Rashotte, C. (1997). *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro–Ed.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific studies of reading*, 11(1), 3-32.

Electronic sources

Hellenic National Corpus (HNC) [http://hnc.ilsp.gr]