

Characteristics of Weak and Strong Readers in a Foreign Language

J. CHARLES ALDERSON

Lancaster University

Linguistics and English Language

Lancaster LA1 4YW

United Kingdom

Email: c.alderson@lancaster.ac.uk

ARI HUHTA

University of Jyväskylä

Centre for Applied Language Studies

Box 35x

Jyväskylä 40014

Finland

Email: ari.huhta@jyu.fi

LEA NIEMINEN

University of Jyväskylä

Agora Center

Box 35

Jyväskylä 40014

Finland

Email: lea.s.m.nieminen@jyu.fi

Author Sequence: Alderson, Huhta, Nieminen

<A>ABSTRACT

The study investigated the cognitive (L1 and FL), linguistic (L1 and FL) and motivational characteristics of weak foreign language (FL) readers in contrast to strong readers in 3 groups of L1 Finnish-speaking learners of English, aged 10, 14, and 17 years. This cross-sectional study covered a wide range of potential correlates, and therefore predictors, of FL reading based on previous research on reading in first, second (L2), and foreign languages. The weakest and strongest FL readers (1 standard deviation below or above the mean reading score) in each age

group were selected for the comparisons reported in this article. The FL (English) skills other than reading were found to be the clearest distinguishing feature between weak and strong FL readers. However, L1 and cognitive skills in L1 and FL, as well as certain dimensions of motivation to use and learn the FL, were also found to separate the two reading groups. The implications of the findings for 2 different hypotheses about the nature of FL reading—the Threshold Hypothesis and the Linguistic Coding Differences Hypothesis—are discussed, as well as the implications for the diagnosis of strengths and weaknesses in reading in an FL.

<END OF ABSTRACT>

Keywords: reading in a foreign language; weak and strong readers; language skills; cognitive skills; motivation; diagnostic assessment

The ability to read in a foreign language (FL; language not spoken in a learner’s environment) is increasingly important in the modern globalised world, yet little is known about how the ability to read in an FL develops, what might underlie weaknesses in FL reading and how to diagnose such weaknesses. This is all the more surprising since the diagnosis of weaknesses in reading in one’s first language (L1) is well established, especially for the more widely spread languages like English, and L1 reading teachers are usually given training in how to diagnose their learners’ weaknesses through pre- and in-service training, and programmes like Reading Recovery (see www.readingrecovery.org; Clay, 1985).

Although the language testing literature frequently refers to diagnostic tests as identifying learners’ strengths and weaknesses, very few truly diagnostic tests exist (see Alderson, Haapakangas, Huhta, Nieminen & Ullakonoja, 2015). Diagnosis of FL learners’ strengths and weaknesses seems to be left up to the classroom teacher, yet in contrast to L1, FL teachers

receive little or no training in how to go about diagnosing such strengths and weaknesses, particularly in FL reading.

<A>LITERATURE REVIEW

Reading in a Foreign Language: A Reading Problem or a Language Problem?

Alderson (1984) raised the question as to whether reading in a FL was a reading problem or a language problem. In a speculative article, he reported many FL teachers as holding the position that their students had problems reading in a FL because they were poor readers in their first language. Alderson concluded tentatively that “although both L1 reading ability and FL language knowledge affect FL reading comprehension, L2 reading appears to be both a language and a reading problem, but with firmer evidence that it is a language problem, for low levels of FL competence, than a reading problem” (cited in Bossers, 1991, p. 48). Alderson speculated that there is likely to be a language threshold beyond which FL readers have to progress before their L1 reading abilities can transfer to the FL – which Clarke (1979) called the ‘short-circuit hypothesis.’

Since then, however, a number of empirical studies have investigated this question, many of which have administered tests of L1 and second or foreign language (SFL) reading comprehension, as well as tests measuring L1 and SFL vocabulary and grammar. Hacquebord (1989) tested 50 L1 Turkish students in basic vocational or basic general secondary education in the Netherlands. Their average age was 13.9 years, and most had immigrated to the Netherlands before the age of 4. Their knowledge of Dutch as a second language (L2; language spoken in the learner’s environment) was four times a stronger predictor of their L2 reading than their L1 reading, at both times of testing (2 years apart). It was claimed that this may have been due to loss of the L1 over time (but no evidence was offered).

Bossers (1991) tested 50 adult native speakers of Turkish, all of whom had passed the highest level of general secondary education in Turkey. They were recent immigrants to the Netherlands and were learning Dutch as a second language in the Netherlands. He found a correlation between L1 Turkish and Dutch as a L2 reading comprehension of .59, but the correlation between L2 comprehension and L2 vocabulary and structures was much stronger at .83. Comparing the low L2 reading group ($n = 35$) with the high L2 reading group ($n = 15$), L2 proficiency was the only significant predictor for the low group, whereas L1 reading was the only significant predictor for the high group. Bossers (1991) concluded that “knowledge of the target language plays a dominant role initially and L1 reading becomes a prominent factor at a more advanced level. This is exactly what a threshold or language ceiling hypothesis would predict“ (p. 56).

Bernhardt and Kamil (1995) found moderate correlations between L1 English and L2 Spanish reading tests and claimed that L1 literacy is a strong predictor of L2 reading at around 20% of variance, but L2 linguistic knowledge consistently accounted for around 30% of the variance. Taillefer (1996) studied L1 French and FL English, and found that L1 reading comprehension and FL knowledge accounted for 35% of the variance on the FL reading test, but FL knowledge accounted for three times more of the variance than did the L1 reading test. Yamashita (2002) tested Japanese students on their L1 Japanese reading, FL English reading and FL English knowledge. Together, L1 reading and FL knowledge accounted for 40% of the variance in FL reading, with FL knowledge being five times stronger than L1 reading.

Despite the differences in L1 tested, status of the SFL, the nature of the tests used (grade point averages, standardised reading and knowledge tests, cloze tests, recall measures, etc.), and the varying sizes and ages of the samples tested, it would appear that the findings consistently

indicate greater importance for SFL knowledge in SFL reading than L1 reading ability, at least until some threshold (undetermined) of SFL proficiency has been passed.

L1 and FL Learning

In fact, the largest influence on SFL reading research has been the L1 research on how children learn to read, and in particular among at-risk readers, rather than dyslexics per se (for an overview of reading development, see Nunes & Bryant, 2009; Pang, 2008; readers can be ‘at risk’ for various reasons such as disability, poor instruction or lack of literacy support at home). Researchers have sought to identify the sorts of learning problems that are experienced by at-risk readers. Vellutino and Scanlon (1986) compared good and poor L1 readers, to find that the poor readers had difficulties with basic L1 language skills, particularly phonology and syntax. A wealth of research since then has confirmed that phonological coding (sometimes known as word decoding) and difficulties with phonological awareness (the conscious ability to recognize, categorize, produce, and manipulate phonological information of the language at the syllable, onset and rime, and phoneme levels) lie at the heart of the problems that children experience when learning to read in their L1, most early research being done in English (Bradley & Bryant, 1983; Liberman & Shankweiler, 1985; Stanovich, 1986a).

Sparks, Ganschow, and Pohlman (1989) proposed that at-risk FL learners also experience a linguistic coding problem when learning the FL. They called their explanation for this problem the Linguistic Coding Deficit Hypothesis (LCDH, later renamed as Linguistic Coding Differences Hypothesis); Sparks and his associates (Sparks & Ganschow, 1993a, 1993b, 1995; Sparks et al., 1998; Sparks et al., 2008) have since published widely on the research justifying this hypothesis.

Numerous studies (reported in Sparks & Ganschow, 1993a) showed that no differences were found in L1 reading comprehension and vocabulary or IQ between at-risk and successful students, but significant differences were found in L1 phonology and syntax, as well as on the Modern Language Aptitude test (MLAT; Carroll & Sapon, 1959). Sparks and Ganschow also reported that methods that directly teach phonology in L1 also result in significant gains on the MLAT. Case studies by the same authors reinforced these findings and the authors argued that cognitive problems relating to basic language skills in L1 are the main cause of problems in learning a FL. They recommended identifying at-risk FL learners by using a battery of L1 measures, including phonology (pseudo-word recognition, spelling, phoneme segmentation), syntax (written grammar), semantics (vocabulary, reading comprehension, oral language), and FL aptitude. Sparks and Ganschow (1993b) argued that poor phonological processing and lack of phonological awareness may give rise to Matthew effects (Stanovich, 1986a, 1986b, 1991), which leads to motivational and attitudinal problems, and not vice versa.

Motivation for FL Learning

MacIntyre and Gardner (1991, 1994) argued that language anxiety is an important cause of individual FL achievement and of differences in FL learning (see also Skehan, 1991).

MacIntyre (1995) asserted that Sparks et al.'s LCDH understates the importance of affective variables; he argued that, among other factors, language anxiety can play a causal role in creating individual differences in language learning.

Sparks and Ganschow (1995) reacted to MacIntyre's criticism by restating their belief that "problems with FL learning are not likely to be primarily the result of low motivation, poor attitude or high anxiety" (p. 235). They pointed out that anxiety can be related to L1 skills as

well as to FL learning, and thus they argued that anxiety cannot be considered on its own, but alongside learners' overall language skills, be they first or foreign language.

However, since the study to be reported here is specifically interested in the diagnosis of strengths and weaknesses in FL reading performance, it is important to explore whether motivation can indeed predict FL reading performance, despite Sparks and Ganschow's 1995 assertions.

More recently, studies of the relationship between motivation and FL proficiency have proliferated (Dörnyei, 2005; Dörnyei & Ushioda, 2011; Iwaniec, 2014; Kormos & Csizér, 2008), and that research has developed and used a wide range of motivational constructs, including Instrumentality, Intrinsic Interest, Motivational Intensity, Parental Encouragement, Anxiety, Self-Regulation, and SFL Self-Concept. Research (e.g., Netten, Droop, & Verhoeven, 2011) has indeed shown that motivation is connected with reading performance in both L1 and L2. Previous research has also shown that motivation is connected with individual differences in second language learning (e.g., Skehan, 1991). Aspects of motivation, including anxiety, are potentially interesting factors for diagnosis because they may not only explain strong vs. weak performance in reading in a SFL, but the educational system can try to do something about them. For example, teachers can try to change learners' negative attitudes and persuade them to see learning in a more positive light and thus become more motivated learners.

Based on recent findings by Kormos and Csizér (2008) it is suggested that learners' age also plays a role in the importance of different motivational constructs. In the study to be reported in this article, we therefore included measures of such motivational constructs as well as the age of learners among the potentially important factors in the difference between weak and strong FL readers.

Methodological Issues

Most of the evidence for the link between L1 language skills and SFL learning is based upon correlational and regression studies of the relationship between various L1 measures, language aptitude (MLAT, Carroll & Sapon, 1959), and SFL achievement, as well as on in-depth case studies. In two studies reported in Sparks et al. (1998), however, the authors examined the differences in L1 skills, FL aptitude, and FL achievement among high school students classified as high-, average-, and low-proficiency FL learners. Because this is a methodology that we partially follow in the current study we now provide details of the Sparks et al. studies.

Both studies reported in Sparks et al. (1998) examined differences in L1 language skills, FL aptitude and FL second-year final course grades. The first study involved 60 females divided into three groups according to their FL proficiency level. High proficiency learners had a mean score at least 1 *SD* or more above the mean ($n = 13$); the average proficiency group ($n = 35$) scored between .99 *SD*'s above the mean to .99 *SD*s below the mean. The low proficiency group ($n = 12$) had a mean score of -1.00 *SD*'s or more below the mean. MANOVA showed overall differences among the three proficiency groups on all L1 language measures and the FL aptitude test. There were significant ($p < .05$) differences between High and Low groups on all L1 and aptitude measures except for phoneme deletion.

Study 2 was similar to Study 1. Thirty-six 10th graders enrolled in second year Spanish, French and German courses took several measures of their L1 language skills, some of which were similar to those in Study 1, plus the MLAT, during their first-year FL course. ANOVAs and post hoc tests showed similar significant differences between the groups as in Study 1. Significant differences between the High and Low groups were found on all the L1 measures (where phoneme deletion was replaced by an informal "Pig Latin" test of phonemic awareness). Sparks et al. (1998) concluded that "students with low average to average (L1) language skills as

reflected in a variety of tests measuring phonology/orthography, semantics, group achievement, and foreign language aptitude may still experience moderate to severe difficulties with foreign language learning in school” (p. 206). There may well be implications for the prediction and diagnosis of potential weaknesses in FL learning: the authors suggested that “foreign language word decoding is a good predictor of overall (oral and written) proficiency in a foreign language” (p. 207).

Although Sparks and associates used standardized measures of L1 cognition and reading and of IQ, they only used FL measures of the same cognitive skills in one study (Sparks et al., 2008), nor were their measures of FL proficiency standardized: They were either course grades, or informal, researcher-made measures somehow linked to the ACTFL descriptors. Nor were reliability coefficients reported in most studies, other than for the researcher-made ACTFL-based proficiency tests.

Meta Analyses

Despite the limitations mentioned before, the studies reported present interesting findings about the importance of L1 and SFL skills in SFL reading, whose tentative conclusions would be strengthened by rigorous meta-analyses.

Melby–Lervåg and Lervåg (2014) conducted a meta-analysis of 82 studies comparing L1 and L2 learners (‘L2’ was defined as the additional language used in the context of English) that looked at reading comprehension and its “underlying components,” which they list as language comprehension, decoding, and phonological awareness. Moderator variables included age, socio-economic status, home language, instructional language, differences between L1 and L2, writing system in L1, consistency of L1 orthography, test type and nonverbal IQ as a moderator to rule out the influence of this factor on group differences. The authors concluded that “unless specific

decoding problems are detected, interventions that aim to ameliorate reading comprehension problems among second language learners should focus on language comprehension skills” (p. 409). Thus, this meta-analysis would appear to support the conclusions of research into Alderson’s 1984 Threshold Hypothesis, namely that reading in a (second/foreign) language is more of a language problem than a reading problem, rather than Sparks et al.’s LCDH. Another conclusion was that there are at least two kinds of struggling readers, those who have decoding problems and those who have comprehension problems, and therefore the prediction and treatment of such different weaknesses in SFL reading might be somewhat different.

Wide-ranging though this meta-analysis is, it still has limitations. A rather low-level or basic set of cognitive skills (primarily decoding) was the focus, rather than higher-order cognitive variables which were not measured, including working memory and automaticity (fluency). Although L1 and L2 vocabulary were measured, other linguistic variables like spelling, syntax, semantics, and writing were not. Nor were reading skills included like inferencing, gist and main idea comprehension, scanning and skimming (so-called expeditious reading), integrating/synthesising and interpreting, reflecting and evaluating.

Jeon and Yamashita (2014) also reported a meta-analysis of L2 reading comprehension and its correlates. The authors argued that a components approach to reading comprehension allows one to observe how different processes interact with each other, as predicted by Stanovich’s (1980) Interactive–Compensatory Model of reading, as well as to assess how much the individual components contribute to the overall skill, be that of reading in L1 or L2. The meta-analysis asked “whether the individual ability difference observed in L2 reading comprehension is better explained by the variance observed in L2 language knowledge (e.g., vocabulary, grammar) or by variance stemming from more language-general, cognitive, or metacognitive processes (e.g., working memory, metacognitive knowledge)” (2014, p. 162). This

meta-analysis is thus relevant both to the claims of Sparks and associates as to the role of L1 knowledge in FL learning, and to the research inspired by Alderson's question whether FL reading poses a reading problem or a language problem.

Ten passage-level SFL reading correlates were included in Jeon and Yamashita's (2014) study, namely SFL decoding, vocabulary knowledge, grammar knowledge, phonological awareness, orthographic knowledge, morphological knowledge, listening comprehension together with L1 reading comprehension, working memory, and metacognition.

The authors concluded that grammar and vocabulary are equally important correlates of SFL reading comprehension. The correlations for SFL language variables were larger than the mean correlations for language-general variables, and the authors argued therefore that SFL reading poses a language problem more than a reading problem. Moreover, SFL reading correlated strongly with SFL listening, providing support for the Simple View of Reading (Hoover & Gough, 1990).

Jeon and Yamashita (2014, p. 196) called for more research on SFL orthographic and morphological knowledge, and cited Jeon (2011) as showing FL morphological knowledge and awareness to contribute substantially to FL reading comprehension, and Kato (2009) as showing that L2 orthographic knowledge was an important contributor to L2 comprehension among higher proficiency readers.

However, Jeon and Yamashita acknowledged similar limitations of their study to those previously described with respect to the Melby-Lervåg and Lervåg (2014) study, namely the limited variety of processing efficiency variables, largely confined to decoding in this study. They also acknowledged the importance of the distinction between L2 and FL contexts.

Summary Discussion

It is clear that weak second or foreign language readers are worse readers than strong SFL readers, but weak at what and why? Are they weak because of L1 reading problems or because their SFL is weak? If they have a reading problem what is that problem, exactly? If they are weaker in their SFL, why, and in which areas do they have weaknesses? In short, what factors underlie their SFL reading weaknesses? Do weak readers have a different profile across a range of factors that distinguishes them from stronger readers? If so, are these factors cognitive, linguistic, affective, motivational, or are they bound up with background features that affect their ability to read in the SFL?

The two major hypotheses that we have now explored—the Threshold Hypothesis and the Linguistic Coding Differences Hypothesis—seem to be in contrast with each other, and related research seems to have reached different conclusions. A number of variables have been explored within the broad areas of cognition, language and affect but within somewhat different frameworks and with somewhat different outcomes. Could it be that insufficient attention has been paid to the range of possible measures that operationalize the researchers' constructs? Might it be that the measures of L1 and/or SFL reading are inadequate in some way, or that L1 and/or SFL language knowledge, skills, and ability have been inadequately construed?

In short, many questions remain to be answered, and if we wish to have a better understanding of what underlies weaknesses in SFL reading in order to develop diagnostic procedures, relevant feedback, and more appropriate interventions in order to help weak readers become stronger readers, then we need to investigate further and more deeply, in order to arrive at clearer answers to at least some of these questions. That is what the exploratory project Diagnosing reading and writing in a second or foreign language (DIALUKI), to be described in the next section, set out to do.

<A>THE DIALUKI PROJECT

DIALUKI was a project that explored the diagnosis of strengths and weaknesses in reading and writing in a SFL. The project comprised a team of applied linguists, language testers, psychologists, and experts in dyslexia whose ultimate aim was to understand what characterises weak SFL readers, and how their reading might best be improved. To that end, we administered a wide range of instruments to explore learners' L1 and SFL skills, cognition and motivation, as described in the Methods section.

The learners came from two different groups: Finnish-speaking learners of English as a FL, and Russian-speaking learners of Finnish as an L2. However, in this article we are concerned with the former group, whilst other publications have addressed the second group (for example, Alderson et al., 2015).

<A>METHODS

Research Questions

Previous studies (e.g., Sparks et al., 1998) and meta-analyses (Jeon & Yamashita, 2014; Melby-Lervåg & Lervåg, 2014) suggested that to improve our understanding of weak FL reading performance we need to investigate FL readers with a wider repertoire of measures. The DIALUKI project sought to do precisely that and therefore the research questions addressed here were the following:

- RQ1. What are the cognitive (L1 and FL), linguistic (L1 and FL), and motivational characteristics of weak readers of English as a FL in contrast to strong readers?
- a. Do weak and strong EFL readers differ in their ability to read in L1 Finnish?
 - b. Do weak and strong EFL readers differ in their English skills other than reading?

c. Do weak and strong EFL readers differ in their cognitive skills?

d. Do weak and strong EFL readers differ in their motivation to use and study English?

RQ2. To what extent are these characteristics similar across three different age groups?

Given the exploratory nature of the study it did not set out to investigate specific hypotheses. However, our general expectation based on previous research was that learners' L1 and FL skills, and possibly also their cognitive performance and motivation, should distinguish weak and strong FL readers, but which of these areas would turn out to be the clearest separators was difficult to predict. As regards the three age groups, we expected to find some differences between them simply because of their different age and amount of prior English studies but the exact nature of these differences could not be predicted, as we were not aware of any previous research comparing age groups in the way we did.

Participants

The participants were three groups of native Finnish speakers learning English as a FL in primary and secondary schools in Finland: 10-year-old children in Grade 4 of primary school, who had been learning English for about 18 months; 14-year-old adolescents in Grade 8 of lower secondary school; 17-year-old students in the second year of academic upper secondary school (gymnasium). These groups were selected since they represent the widest possible range of ages and proficiency levels in English available for systematic study in the Finnish educational system: The 4th graders had just enough prior studies to make it meaningful to investigate literacy skills in English and the second year in the gymnasium is the last stage at which students preparing for the high-stakes final examination (typically in their third year) could be expected to participate in such school-external activities as the study reported here. The English test results

indicate that we did indeed manage to cover a wide range of proficiencies. The 8th graders and gymnasium students took the DIALANG English reading tests (see Appendix A); therefore, we have estimates of their Common European Framework of Reference (CEFR) levels in reading. The most typical level for the 8th graders was A2 and for the gymnasium students B1 or above, even up to C1/C2. The 4th graders' English reading scores cannot be reliably related to the CEFR but presumably most of them were at A1 or below. The CEFR-based ratings of participants' writing ability in English yielded very similar results. The study reported here is a cross-sectional study across these three groups. The total number of participants involved was 203 4th graders (girls 52%, boys 48%); 204 8th graders (girls 56%, boys 44%); 205 gymnasium students (girls 53%, boys 47%).

Instruments

The independent variables were the categorizations of FL readers into weak and strong readers based on their performance on standardized English reading tests (see *Analyses*). The 4th graders were given the Pearson Test of English (PTE) Young Learners Test at levels A1/A2 on the Common European Framework of Reference (CEFR; Council of Europe, 2001). The 8th graders were given the PTE General (Reading) at CEFR levels A2/B1 as well as the DIALANG Test of Reading (Intermediate) at levels B1/B2. The gymnasium students were given the PTE General Reading Test and the DIALANG Reading Test (Advanced) at a higher level (B1–C1 on the CEFR) than the tests taken by the 8th graders (see Appendix B for the descriptive statistics and reliabilities).

A number of cognitive (L1 and FL), linguistic (L1 and FL), and motivational measures as dependent variables were also administered within two months. The 17 cognitive measures used covered phonological processing, working memory and lexical retrieval; both L1 and FL

versions of the measures were administered. A total of five L1 and four FL measures were also taken by the learners (e.g., vocabulary, writing and segmentation tasks). In addition, they responded to an extensive questionnaire covering seven dimensions of motivation on using and learning English (e.g., intrinsic and instrumental motivation, anxiety, self-regulation). For details of the instruments see Appendices A to C.

Analyses

To compare learners' reading proficiency in English with the various cognitive, linguistic, and motivational variables, learners' English reading scores (in fact, Rasch measures since they provide more valid indicators of ability than raw scores; see, e.g., Engelhard, 2013) were first converted to z-scores and then the learners were divided into three groups: strong, medium, and weak readers. We used one standard deviation above/below the mean as the cut-off points (the same procedure was used by, e.g., Sparks et al., 1998). In the current study, however, only the weak and strong FL readers are investigated. The number of weak and strong readers varies to some extent between different analyses, for example, due to absence of some students on the day particular measures were administered. In most analyses, there were 15–22 weak FL readers and 20–30 strong FL readers (see Appendices D–G for details).

Before comparing weak and strong FL readers, we condensed the 17 different cognitive measures into a more manageable number of constructs. The other measures had either already been turned into factors (motivation) or were large enough units in their original form (e.g., reading, writing, vocabulary). Since the 8th graders and gymnasium students took exactly the same cognitive measures, a joint factor analysis could be performed for them. However, the 4th grade data had to be analysed separately because some of the measures were shortened or were not administered at all in order to lessen the burden caused by the study to these young learners.

Factor analyses of the cognitive measures yielded five factors for the 4th graders and four factors for the two older age groups. The decision on the number of factors was based on examining the eigenvalues, scree plots, and the interpretability of the factors. The factor content was remarkably similar, but not identical, across the groups. In the two older groups, the biggest factor featured rapid naming and reading of individual words, in both L1 and FL, but in the 4th grade group the key components of this factor were the rapid word recognition tasks, again in both languages, even though the rapid word list reading tasks also loaded heavily on this factor (see Appendix H). Most of these loadings were at the .7 or .8 level. In some cases the three types of measures which aimed at tapping automatization and speed of access to vocabulary (i.e., rapidly presented words, word list reading, and rapid alternating stimulus naming) also loaded on other factors than the fluent reading factor(s). In grade 4, the Finnish word list reading task also had a lot in common (over .5 loading) with the Finnish phonological skills factor. In the two older groups both the word list and alternating stimulus tasks, in both L1 and FL, also loaded on the phonological factor, and in particular the English versions of the word list and alternating stimulus tasks also loaded on the working memory factor. Thus, there is a clear overlap between these and the other factors, which might suggest that there is a higher order cognitive factor in the background. However, since our aim was to obtain more detailed, potentially diagnostic information about the distinguishing characteristics of weak and strong FL readers, a multi-factorial approach was used.

In both weak and strong reader groups it was possible to identify a working memory factor in which the core comprised the two backwards digit span tasks. As might be expected, working memory plays a significant role in many other cognitive tasks and, consequently, several other cognitive measures loaded at least moderately (typically between .4 and .5) on that factor.

Two phonological factors were also identified in both groups. One was based on specific Finnish phonological tasks, whereas the other differed: For the older students, the factor related mainly to English phonological tasks and for the 4th graders it was based on phonological tasks in both languages.

Overall, the first, ‘fluent reading’ factor explained 26% of variance in the L1 and FL cognitive measures in both the combined 8th grade and gymnasium group and in the 4th grade group, whereas the second factor (i.e., mainly English phonology in the two older groups or another ‘fluent reading’ factor in the 4th grade) explained only 6–8% of the variance.

To examine the differences between the strong and weak FL readers statistically, a number of MANOVAs were run on groups of independent variables (e.g., L1 linguistic measures, FL linguistic measures, cognitive factors and motivation) to investigate any overall differences between the weak and strong FL readers in each group of variables of interest. These were followed by univariate tests (in MANOVA) to examine the differences in more detail.

<A>RESULTS

The results are presented mainly for Research Question 1; RQ2 (comparing the results across the age groups) is partly covered in this section and partly in Discussion.

RQ1a: Do weak and strong EFL readers differ in their ability to read in L1 Finnish or their L1 Finnish skills in general?

To address RQ1a, the following L1 Finnish tests were used as dependent variables in MANOVA analyses: reading, writing, vocabulary, and segmentation (time and accuracy). For the two older groups spelling error detection tests were also available (speed and accuracy). The weak vs. strong EFL reading group was the independent variable. The overall multivariate tests

comparing the weak and strong readers were highly significant ($p < .001$) in all three age groups, as shown in Appendix D.

Univariate tests of the between-subjects effects of the L1 measures were computed to investigate which first language skills most clearly separated weak and strong EFL readers. The significantly distinguishing L1 skills in the three age groups were: (a) 4th grade: Reading, Writing, Segmentation accuracy, (b) 8th grade: Reading, Vocabulary size, Spelling error detection accuracy, Spelling error detection speed, Segmentation speed, and (c) gymnasium: Vocabulary size, Segmentation speed, Reading, Spelling error detection speed, Spelling error detection accuracy, Segmentation accuracy.

Several aspects of L1 proficiency distinguished weak and strong EFL readers. The best separators differed somewhat across the groups but in all cases an ability to read in one's L1 (Finnish) was among them. For the two younger groups, L1 reading skill was the best separator of the two FL reading groups with a large effect size (Cohen's $d = 1.38$ and 1.48 , respectively; see Appendix D). It was also an important factor in the gymnasium group ($d = .98$). Writing in L1 was significant only for the youngest group ($d = 1.32$). Knowledge of L1 words distinguished weak and strong FL readers quite well among the two older age groups ($d = 1.21$ for the 8th graders and 1.15 for gymnasium).

The L1 segmentation task, which combines knowledge of vocabulary, morphology, and syntax, had an interesting relationship with weak vs. strong FL reading. Segmentation accuracy was a useful separator ($d = 1.07$) for the young 4th graders and also for the gymnasium group ($d = .75$) but not for the 8th graders. Speed of segmentation distinguished the reading groups in the gymnasium very well ($d = 1.55$) and to some extent in the 8th grade ($d = .59$), but not in the 4th grade.

The L1 spelling error detection task consisted of a list of unrelated words, each containing one error; the learners' task was to indicate where each word had an error but not to correct it. The speed measure was the number of errors identified in the time given (3½ minutes). The accuracy measure was the proportion of correctly indicated errors. Spelling error detection functioned moderately well in both the age groups it was administered to as it distinguished the weak FL readers from the strong (d ranged from .70 to 1.02).

RQ1b: Do weak and strong EFL readers differ in their English skills other than reading?

The following FL English tests were used as dependent variables in MANOVA analyses: writing, dictation, vocabulary, segmentation (time & accuracy). The independent variable was the strong vs. weak EFL reading group. The overall multivariate tests comparing the weak and strong FL readers were highly significant ($p < .001$) in all three age groups (Appendix E).

Univariate analyses of the FL measures shed more light on which FL skills most clearly separated weak and strong EFL readers. In summary, the distinguishing FL skills were: (a) 4th grade: Writing, Dictation, Vocabulary size, Segmentation accuracy, (b) 8th grade: Writing, Dictation, Vocabulary size, Segmentation accuracy, Segmentation speed, and (c) gymnasium: Writing, Dictation, Vocabulary size, Segmentation accuracy, Segmentation speed.

Several aspects of English proficiency significantly distinguished weak and strong EFL readers. The best overall separator was vocabulary size with very large effect sizes in all three groups ($d = 2.33, 3.86, \text{ and } 5.01$, from the youngest to the oldest group). Writing and dictation performance also clearly separated the weak and strong FL readers: The effect sizes for writing for the three age groups were $d = 1.39, 3.92, \text{ and } 3.27$, and for dictation $d = 2.13, 3.67, \text{ and } 5.14$, respectively. Segmentation accuracy also separated the two FL reading groups across all age levels, with large effect sizes ($d = 1.88, 2.72, \text{ and } 2.67$). Interestingly, speed of segmentation did

not distinguish the readers in the 4th grade but in the 8th grade and gymnasium it clearly did ($d = .71$ and $.93$, respectively).

RQ1c: Do weak and strong EFL readers differ in their cognitive skills?

The cognitive factors identified in the factor analyses were used as dependent variables in MANOVA analyses and the weak and strong EFL reading groups were the independent variable. The overall multivariate tests were highly significant in all age groups (Appendix F).

Univariate analyses of the cognitive factors in both L1 and FL shed more light on which aspects of cognitive performance most clearly separated weak and strong readers. To summarise, the distinguishing cognitive features in the three age groups were: (a) 4th graders: Fluent reading 1 (based on word level reading and naming tasks), Fluent reading 2 (based on rapid recognition of words), General phonological skills, (b) 8th graders: Working memory, Phonological skills in English, Fluent reading, and (c) gymnasium: Working memory, Phonological skills in English, Fluent reading, Phonological skills in Finnish.

Several cognitive factors distinguished weak and strong EFL readers but somewhat differently at different ages. For the 10-year-old 4th graders the only clearly distinguishing factor was the ability to read and recognise words rapidly both in L1 and FL. In contrast, other aspects of cognition were more important in separating the readers in the older age groups, namely working memory (based on both L1 and FL tasks) and phonological processing ability in English. Fluent word-level reading (in both L1 and FL) played some part in the older groups, but was not as dominant as among the 4th graders.

Judging by the effect sizes, almost all the statistically significant differences were also large. The best separator of weak and strong readers in the 8th grade and gymnasium was working memory ($d = 1.74$ and 1.69 , respectively), followed closely by English phonology ($d =$

1.56 and 1.53). Fluency in word level reading in L1 and FL was also a fairly clear separator across all groups ($d = 1.19, 1.34,$ and $.98,$ in the 4th, 8th, and gymnasium, respectively).

RQ1d: Do weak and strong EFL readers differ in their motivation to use and study English?

The motivational factors that had been identified earlier in the project were used as dependent variables in MANOVA analyses: Anxiety, English FL self-concept, Instrumental motivation, Intrinsic motivation, Motivational intensity, Parental encouragement, and Self-regulation.

The overall multivariate tests comparing the weak and strong EFL readers were highly significant in the two older age groups, but not quite ($p = .060$) significant for the 4th graders (Appendix G).

Univariate tests were computed to identify which motivation factors separated weak and strong readers, including the 4th graders, even though their overall test did not quite reach significance, in order to check whether any of the dimensions were relevant separators. The distinguishing motivational factors were: (a) 4th graders: English FL self-concept, (b) 8th graders: English FL self-concept, Instrumental motivation, Intrinsic interest, and (c) gymnasium: English FL self-concept, Intrinsic interest, Anxiety, Motivational intensity.

The most important motivational factor that distinguished weak and strong FL readers was learners' concept about themselves as learners of English. It was the only significant separator for the 4th graders ($d = 1.03$), and its importance increased with age and experience in learning the FL: The effect size rose to $d = 1.97$ in the 8th grade and further to $d = 3.09$ in the gymnasium. Instrumental motivation was a separating factor for the 8th graders ($d = 1.00$) but not for the other groups. In contrast, intrinsic motivation was important for the gymnasium students

($d = 1.61$) who were also the only group in which anxiety ($d = 1.24$) and motivational intensity ($d = .94$) separated the two reading groups.

Some findings pertaining to Research Question 2 about age differences were already presented but they will be covered in the discussion more systematically.

<A>DISCUSSION

The goal of this study was to investigate the cognitive (L1 and FL), linguistic (L1 and FL), and motivational characteristics of weak English as a foreign language readers in contrast to strong readers. The study covered an unusually wide range of characteristics and used standardized instruments to measure most of these. We found that the FL (English) skills other than reading were the clearest distinguishing feature between weak and strong EFL readers at all age levels. Vocabulary size and the ability to write, to do dictation tasks (which involves listening and writing), and to segment the FL also separated the weak and strong readers. However, first language skills, particularly reading and vocabulary knowledge, but also segmentation ability, were also fairly good separators. Furthermore, most but not all of the basic cognitive skill factors in L1 and FL, especially working memory capacity, phonological skills and fluent word level reading, separated the two EFL reading groups. The most important motivational factor that distinguished the EFL readers across the three age groups was learners' concept of themselves as learners of English.

Trends and Differences Across Age Groups

It should be borne in mind that the three age groups represent three broadly different proficiency groups as the 4th, 8th, and gymnasium students had studied English for 1½, 5½, and 8½ years, respectively. As mentioned earlier, the 4th graders were typically at the CEFR level A1

or below, the 8th graders at A2, and the gymnasium learners at B1 or above as far as their English reading level is concerned.

Two broad similarities were evident across the three age groups. First, the content of the motivational and L1 and FL cognitive factors was rather similar in all groups. Second, the relative importance of FL, L1, cognitive and motivational factors in distinguishing weak and strong EFL readers was the same: FL skills were always clearly the best separators.

Despite these two similarities, some differences across groups may be important diagnostically as they imply that certain indicators may be important at certain ages or stages of learning. It is worth remembering that the 4th and 8th graders represent whole age groups whereas the gymnasium is typically attended by the more academically oriented students.

Even though command of the FL skills other than reading was the most important separator of weak and strong FL readers in all age groups, their importance increased with learners' age and proficiency (see Appendix E). FL vocabulary knowledge is a good example: Its effect size increased from $d = 2.33$ in grade 4 to $d = 3.86$ and $d = 5.01$ in the two older groups. The same trend could be observed for FL writing, and, less clearly, for dictation. Segmentation accuracy was also a clear separator (with $d = 1.88$ – 2.72) for all ages. These findings do not suggest a threshold such that L1 reading becomes more important with age or increasing FL proficiency and thus might appear to contradict Bossers (1991). However, he investigated L2 rather than FL learners.

Interestingly, segmentation speed did not separate weak readers from strong in the youngest group but did so for the 8th grade and gymnasium (with $d = .71$ and $.93$, respectively). This suggests that the speed of performing a task that involves both vocabulary and syntactic knowledge is potentially a useful indicator of reading (comprehension) problems for learners who have studied the language for some considerable time.

Among first language skills, L1 reading was the only consistently good separator of weak and strong FL readers (with $d \geq .98$). L1 writing separated weak from strong FL readers in the two younger groups only and L1 vocabulary knowledge only in the two older groups. The latter finding is particularly interesting, as the vocabulary test was very easy (see Appendix C): learners who failed to achieve the maximum score on L1 vocabulary were likely to be among the weakest FL readers. Another diagnostically interesting finding was that both L1 segmentation and L1 spelling error detection often distinguished weak and strong FL readers (Appendix D). Spelling error detection was administered only in the two older groups but both the speed and accuracy in detecting the errors separated the two reading groups. Segmentation accuracy did the same in both the youngest and oldest age group but was only a marginal separator in the middle, 8th grade group. Noteworthy was the large effect size ($d = 1.55$) of L1 segmentation speed in the gymnasium group; it appears that problems in rapid segmentation of either L1 or FL text may be an indicator of FL reading comprehension problems even among learners with considerable experience of study of the FL.

Cognitive factors tended to separate weak and strong readers less clearly in the youngest (4th grade) group (with only one $d > 1.00$). In contrast, three cognitive factors distinguished the weak and strong readers very clearly (mostly $d > 1.00$) in the two older groups. Firstly, fluent word level reading (based on both L1 and FL tasks) separated FL readers quite well among the two younger groups but less well in the gymnasium. Secondly, in contrast, the working memory factor (based on both L1 and FL tasks) only distinguished weak from strong in the two older groups. Thirdly, a factor based on phonological tasks presented in English also separated weak and strong readers clearly in the two older groups, which reinforces the trend noticed earlier that FL skills in general are increasingly important as separators of weak vs. strong FL readers as learners' age and experience with the language increases.

As for motivation, although self-concept distinguished weak from strong in all three age groups, as learners' age increased, more and more motivational variables separated weak readers from strong. Thus we might tentatively conclude that with increasing age, motivation to learn or read the FL decreases among the weakest readers, or that the gap between the strong and weak readers seems to become wider with respect to motivation. Alternatively, a lack of improvement in one's reading ability may bring about decreasing motivation. However, it was somewhat encouraging to find that some aspects of motivation that can potentially be affected by teachers and parents, such as intrinsic motivation and motivational intensity, did separate weak and strong FL readers, particularly among the gymnasium learners.

Moreover, as one reviewer of this study pointed out, while FL self-concept was found to be a significant separator in all age groups, anxiety was significant only for the gymnasium group, which is the oldest and the most proficient. This finding is plausibly related to the fact that our gymnasium learners in the second year of upper secondary school were closer than the younger learners to taking their high-stakes school-leaving examinations.

Theoretical Implications

Previous studies have led to several different hypotheses about the nature of problems reading an SFL. One hypothesis is that FL reading is more of a reading problem (including weakness in L1 reading) than a FL language problem. Conversely, problems in FL reading may be simply a problem of FL knowledge. Earlier we considered two other hypotheses, – the Threshold Hypothesis (Alderson, 1984) and the Linguistic Coding Differences Hypothesis (Sparks et al., 1989). According to our findings, FL knowledge was the best separator of weak and strong FL readers. Much bigger effect sizes were discovered for the FL skills, sometimes exceeding $d = 5.00$ (see Appendix E), than for the L1 and cognitive skills. These findings, then,

support the Threshold Hypothesis more than the LCDH and are in line with the conclusion of the meta-analysis by Melby–Lervåg and Lervåg (2014) that training L2 skills is the best treatment for most learners with L2 reading problems (even if their operationalization of L2 skills is oral vocabulary and oral comprehension more generally).

However, our study demonstrated that certain aspects of cognition (usually based both on L1 and FL tasks) and L1 skills also separated weak and strong FL readers, albeit often only with moderate effect sizes. As reported earlier, some studies reported in, for example, Sparks & Ganschow (1993a) suggest that differences in L1 cognitive and linguistic skills can explain differences in FL skills to some extent. Research on dyslexia (e.g., Helland & Morken, 2015) has demonstrated how important basic cognitive skills are in L1 reading.

Melby–Lervåg and Lervåg (2014) suggested that their recommendation of focusing on improving weak L2 readers' proficiency in L2 may not apply to learners with specific decoding problems. It may be, therefore, that the two theories (the Threshold Hypothesis and the Linguistic Coding Differences Hypothesis) apply to somewhat different kinds of learners. It might be that the Threshold Hypothesis is most appropriate in explaining FL reading among the general population of learners, most of whom are without specific reading problems, as was the case in our study, and that the Linguistic Coding Deficit Hypothesis is most appropriate for at-risk learners. A full investigation of this would probably require the inclusion of a fairly large number of at-risk L1 readers and a comprehensive measurement of their cognitive, L1 and FL skills.

However, there are several caveats to this speculation. One is that previous research has looked at L2 learners, where learners are learning their second language in contexts where they are surrounded by the L2, which very likely facilitates the learning of that second language. Our study, however, has investigated the issue in an FL setting where learners are not surrounded by

the FL, even if the FL in question (English) has some presence in many youngsters' lives in a country such as Finland.

Secondly, previous studies have not focused on age differences, whereas we have. Furthermore, our learners were arguably still developing their cognitive skills in their L1 and in their FL. Moreover, although we have discovered that FL weak readers are clearly weaker in a range of FL skills than stronger readers, in other words that FL reading problems are more of an FL proficiency problem than a problem of L1 reading abilities, we do not know whether even the stronger readers have reached a threshold of FL proficiency that might allow their L1 reading abilities (or strategies) to transfer to their FL reading.

Importantly, we have reported on a cross-sectional study, not a longitudinal one, although we have also conducted a longitudinal study, which many earlier SFL reading researchers have not. We may eventually find evidence of individual development over time that reveals an FL threshold beyond which L1 reading abilities can transfer to FL reading. But in this study, we cannot claim that a threshold has been reached that is associated with increasing age.

>B>Implications for Further Research

Specific implications from our results for further research on diagnosing SFL reading are that the instruments used in research, the characteristics of the learners, and the context in which learning is taking place should be elaborated more carefully than has usually been the case. The nature and quality of the data gathering instruments need serious attention.

First, the FL reading construct should be defined and operationalised carefully, since diagnosing reading for synthesising information, for example, may be different from diagnosing skimming and scanning or inferencing or other component skills. Specifically we have not referred to or used a number of theories of comprehension developed in L1 reading research, like

that of Stanovich's Interactive-Compensatory Model (1980), or Kintsch's ideas on discourse processing and his Construction-Integration Model of text comprehension (2004), or the role of background knowledge and long-term working memory in comprehension.

Second, the selection and operationalization of the predictors is important. Previous research, including ours, has covered a range of potential predictors but there are still gaps such as orthographic and morphological knowledge of the SFL (see Aro & Wimmer, 2003; Jeon & Yamashita, 2014). How particular predictors are operationalised also matters. For example, vocabulary size is already a well-known predictor of SFL reading ability, but it remains to be seen whether improving weak readers' depth or diversity of vocabulary and vocabulary learning strategies, for instance, will increase their reading ability.

In addition, the language through which the instruments are presented is an important consideration as it is very difficult to find tasks that do not involve any degree of language knowledge.

A broad enough range of both dependent and independent variables should be used in diagnostic research, unless there are good reasons, based on theory and previous studies, to limit their range. Even our study with its wide coverage of skills was limited, for practical reasons, in what was measured compared with what could ideally have been measured.

A related study looking at the relation between learners' background characteristics and SFL reading (Huhta, Alderson, Nieminen & Ullakonoja, 2016) has suggested that very few background variables are able to predict performance in SFL reading, but it may be that some such variables are useful diagnostically. Particularly interesting is the amount of SFL reading or other use of the SFL in individual learners' free time. L1 reading research has found that primary school students' reading frequency was the best predictor not only of performance on L1 reading tests but also the student's growth as a reader from the second to the fifth grade (Anderson,

Wilson, & Fielding, 1998). Since free-time use of an SFL is clearly variable, it may be a useful factor to manipulate when attempting to improve learners' SFL reading skills. More generally, Huhta et al. (2016) point out the need to consider the characteristics of the learner and the context more carefully, as the diagnosis—and possible treatment—of learners of different ages, proficiency levels and contexts may differ. In particular, we need research in different L2 and FL contexts and across different L1 and SFL pairs than were studied in this research. Finnish has a transparent orthography and English an opaque orthography (cf. Jeon & Yamashita's, 2014, previously mentioned point). Different L1/SFL pairs might well give different results.

Finally, the current study was cross-sectional, and while we argue that it has shed new light on the role of different factors that are characteristic of weak FL readers, in contrast to strong readers, longitudinal and intervention studies are also needed to confirm and complement the present results on how development in FL reading relates to cognitive, linguistic, and motivational characteristics and which of such characteristics are important simply for predicting or identifying problems in development and which are also useful for treatment purposes.

ACKNOWLEDGMENTS

This study was funded by the Academy of Finland, the University of Jyväskylä, the UK Economic and Social Research Council (ESRC), and Pearson Language Tests.

REFERENCES

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 1–24). London: Longman.

- Alderson, J. C., Haapakangas, E.-L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2015). *The Diagnosis of Reading in a Second or Foreign Language*. New York / London: Routledge.
- Anderson, R., Wilson, P., & Fielding, L. (1998). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 23, 285–303.
- Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, 24, 621–635.
- Bernhardt, E., & Kamil, M. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the Linguistic Threshold and the Linguistic Interdependence Hypotheses. *Applied Linguistics*, 16, 15–34.
- Bossers, B. (1991). On thresholds, ceilings and short-circuits: The relation between L1 reading, L2 reading and L2 knowledge. *AILA Review*, 8, 45–60.
- Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, 301, 419–421.
- Carroll, J., & Sapon, S. (1959). *Modern Language Aptitude Test (MLAT): Manual*. San Antonio, TX: Psychological Corporation.
- Clarke, M. (1979). Reading in Spanish and English: Evidence from ESL students. *Language Learning*, 29, 121–150.
- Clay, M. (1985). *Reading: The patterning of complex behavior* (3rd ed.). Exeter, NH: Heinemann.
- Council of Europe (2001). *The Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. London: Longman Pearson.

- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation*. Harlow, UK: Pearson Education.
- Engelhard, G. (2013). *Invariant measurement. Using Rasch models in the social, behavioral, and health sciences*. New York/London: Routledge.
- Hacquebord, H. (1989). *Tekstbegrip van Turkse en Nederlandse leerlingen in het voortgezet onderwijs*. [Reading comprehension of Turkish and Dutch students attending secondary schools.] Groningen, The Netherlands: RUG. Cited in Bossers (1991).
- Helland, T., & Morken, F. (2015). Neurocognitive development and predictors of L1 and L2 literacy skills in dyslexia: A longitudinal study of children 5–11 years old. *Dyslexia* (online, Oct 2015) DOI: 10.1002/dys.1515
- Holopainen, L., Kairaluoto, L., Nevala, J., Ahonen, T., & Aro, M. (2004). *Lukivaikeuksien seulontamenetelmä nuorille ja aikuisille* [Screening procedure of reading difficulties for adolescents and adults]. Jyväskylä, Finland: Niilo Mäki Institute.
- Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160.
- Huhta, A., Alderson, J. C., Nieminen, L., & Ullakonoja, R. (2016). The role of background factors in the diagnosis of SFL reading ability. In D. Tsagari & J. V. Banerjee (Eds.), *Contemporary Second Language Assessment* (pp. 125–146). London: Bloomsbury.
- Iwaniec, J. (2014). Motivation of pupils from southern Poland to learn English. *System*, 45, 67–78.
- Jeon, E. (2011). Contribution of morphological awareness to second-language reading comprehension. *Modern Language Journal*, 95, 217–252.
- Jeon, E., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160–212.

- Kato, S. (2009). Suppressing inner speech in ESL reading: Implications for developmental changes in second language word recognition processes. *Modern Language Journal*, *93*, 471–488.
- Kintsch, W. (2004) The Construction-Integration Model of text comprehension and its implications for instruction. In R. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed.) (pp. 1270–1328). Newark, DE: International Reading Association.
- Kormos, J., & Csizér, K. (2008). Age-related differences in the motivation of learning English as a foreign language: Attitudes, selves and motivated learning behavior. *Language Learning*, *58*, 327–355.
- Lieberman, I., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education*, *6*, 8–17.
- Lindeman, J. (2005). *Ala-asteen Lukutesti ALLU* [Reading test for primary school] (3rd ed.). Turku, Finland: Centre for Learning Research.
- MacIntyre, P. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *Modern Language Journal*, *79*, 90–99.
- MacIntyre, P., & Gardner, R. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language Learning*, *41*, 513–534.
- MacIntyre, P., & Gardner, R. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, *44*, 283–305.
- Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*, 400–433.

- Netten, A., Droop, M., & Verhoeven L. (2011). Predictors of reading literacy for first and second language learners. *Reading and Writing, 24*, 413–425.
- Nunes, T., & Bryant, P. (2009). *Children's reading and spelling. Beyond the first steps*. Malden, MA/Oxford, UK: Wiley–Blackwell.
- Pang, J. 2008. Research on good and poor reader characteristics: Implications for L2 reading research in China. *Reading in a Foreign Language, 20*, 1–18.
- Reading Recovery. Accessed 30 January 2016 at www.readingrecovery.org
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of the two new versions of the Vocabulary Levels Test. *Language Testing, 18*, 55–88.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13*, 275–298.
- Sparks, R., Artzer, M., Ganschow, L., Siebenhar, D., Plageman, M., & Patton, J. (1998). Differences in native-language skills, foreign-language aptitude, and foreign-language grades and high-, average-, and low-proficiency foreign-language learners: two studies. *Language Testing, 15*, 181–216.
- Sparks, R., & Ganschow, L. (1993a). The impact of native language learning problems on foreign language learning: Case study illustrations of the Linguistic Coding Deficit Hypothesis. *Modern Language Journal, 77*, 58–74.
- Sparks, R., & Ganschow, L. (1993b). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. *Modern Language Journal, 77*, 289–302.
- Sparks, R., & Ganschow, L. (1995). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *Modern Language Journal, 79*, 235–244.

- Sparks, R., Ganschow, L., & Pohlman J. (1989). Linguistic coding deficits in foreign language learners. *Annals of Dyslexia*, 39, 179–195.
- Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *Journal of Educational Psychology*, 100, 162–174.
- Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 16, 32–71.
- Stanovich, K. (1986a). Explaining the variance in reading ability in terms of psychological processes: What have we learned? *Annals of Dyslexia*, 36, 67–96.
- Stanovich, K. (1986b). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.
- Stanovich, K. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*, 26, 7–29.
- Taillefer, G. (1996). L2 reading ability: Further insight into the Short-Circuit Hypothesis. *Modern Language Journal*, 80, 461–477.
- Vellutino, F., & Scanlon, D. (1986). Linguistic coding and metalinguistic awareness: Their relationship to verbal and code acquisition in poor and normal readers. In D. Yaden & S. Templeton (Eds.), *Metalinguistic awareness and beginning literacy: Conceptualizing what it means to read and write* (pp. 115–141). Portsmouth, NH: Heinemann.
- Yamashita, J. (2002). Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading*, 25, 81–95.

APPENDIX A

Measures Used in the Study

A more detailed description of the measures can be found at www.jyu.fi/dialuki (select In English and then Instruments).

a) The Dependent Variables (Foreign Language Reading Tests)

Reading in FL	
4 th graders	Pearson Test of English (PTE) Young Learners Test Two tasks, 12 items targeting CEFR levels A1–A2.
8 th graders	Pearson Test of English General Five tasks, 25 items targeting A2–B1. DIALANG English Reading 30 items based on over 20 texts, targeting A2–B2.
Gymnasium	Pearson Test of English General Five tasks, 25 items targeting B1–B2. DIALANG English Reading 30 items, based on over 20 texts, targeting B1–C1/C2.

b) The Independent Cognitive Measures

Almost all tasks were delivered mostly on computer and separately in L1 and FL.

Working Memory

Backwards digit span	Number of items (k) = for L1 14 items (or 7 digit levels), for SFL 10 items (5 digit levels)
----------------------	--

Automatisation and Speed of Lexical Access

(Fluent Word Level Reading)

Rapidly presented words	Recognizing single words that flash on the screen for 80 milliseconds and are then masked with non-letter characters (e.g., #&:α). $k = 14$ (L1), $k = 12$ (FL)
Rapid word list reading	Reading aloud as many words as possible in one minute from a list of 105 words.
Rapid alternating stimulus (RAS) naming	Naming colours, numbers and letters (or objects), mixed in random order in a matrix of 10 items per row. $k = 50$ ($k = 30$ in the FL version for the 4 th graders)

Phonological Processing/Awareness

Pseudo-word reading	Reading aloud pseudo-words presented one by one on the screen. $k = 10$
Pseudo-word spelling (only in L1)	Spelling pseudo-words, each presented orally twice. Each item was presented twice. $k = 12$
Pseudo-word repetition	Repeating orally presented pseudo-words. $k = 10$
Phoneme deletion	Repeating an orally presented pseudo-word: first as it is and then

without a specific sound. $k = 12$ (L1), $k = 8$ (FL)

Common unit Identifying (by saying aloud) the sound that is common in a pair of orally presented pseudo-words. $k = 10$

Note. The pseudo-word reading in FL and the common unit in FL were not administered to the 4th graders.

c) The Independent Linguistic Measures

L1 Vocabulary Size	DIALANG Vocabulary Size Placement Test Marking familiar/known words in a list of real and pseudo-words. $k = 75$
FL Vocabulary Size	Schmitt, Schmitt and Clapham Vocabulary Levels Test (Schmitt et al., 2001) $k = 60$ (4 th grade), 90 (8 th grade), 120 (gymnasium)
Segmentation in L1 and FL	Dividing a short text with no punctuation or word boundaries into separate words (e.g., thepigsweresohappytheysangthissong → the pigs were so happy they sang this song). L1: $k = 36 / 73 / 73$ words (4 th grade, 8 th grade, gymnasium); FL: $k = 51 / 59 / 71$ words (4 th grade, 8 th grade, gymnasium)
Detecting spelling errors (L1 only)	Indicating the location of spelling error in each word in a list of discrete words (Holopainen et al., 2004). Time allowed: 3½ minutes. $k = 100$
Dictation (FL only)	Writing down a recorded passage which was first presented as a

	whole and then in smaller units.
Reading in L1	4 th grade: a text with 12 multiple choice questions from the standardised Finnish Reading Comprehension Test ALLU (Lindeman, 2005). 8 th grade and gymnasium: three passages with 11 items from the PISA test in Finnish.
Writing in L1 and FL	CEFLING and PTE General 4 th grade: an opinion task in L1 and FL from CEFLING (a previous study). 8 th grade and gymnasium: two L1 CEFLING tasks (an opinion and a message), one FL task from CEFLING and two tasks from the PTE General (opinion, partly narrative).

Note. The spelling error detection task was not administered to the 4th graders.

d) The Independent Measures of Motivation and Anxiety

Iwaniec Motivation Questionnaire*

Adapted from Iwaniec (2014).

The participants answered on a 5-point Likert scale (1 = not true at all; 5 = absolutely true).

The order of items was randomised.

Anxiety

$K = 6$

Examples

I tremble when I know I'm going to be called on in English class.

	Even if I'm well prepared for the English class, I feel anxious about it.
English FL Self-	$K = 6$
concept	Examples: Compared to other students, I'm good at English I have always done well in English.
Instrumentality	$K = 6$
	Examples: The things I want to do in the future require me to use English I study English because it will be useful to get a job
Intrinsic Interest	$K = 6$
	Examples: I am curious about how people communicate in English. I find learning English enjoyable.
Motivational	$K = 5$
Intensity	Examples: When studying English, I try to do it with my best effort When I get my assignments back from my English teacher, I check them and try to learn something from my mistakes.
Parental	$K = 7$
Encouragement	Examples: My parents consider English an important subject My parents encourage me to study English also in my free time
Self-regulation	$K = 6$

Examples:

I try to find opportunities to practice my English

I have my own ways of studying English vocabulary

*These measures were administered in L1 Finnish, but the items addressed motivation for FL English. The wording of some of the statements was somewhat simplified in the 4th graders' motivation questionnaire.

APPENDIX B

Means, Standard Deviations, and Reliabilities of the English Reading Tests

Reading Measures			
(raw scores) *			
	4th Grade	8th Grade	Gymnasium
	<i>M / SD</i>	<i>M / SD</i>	<i>M / SD</i>
Pearson Young Learners Test of English (<i>k</i> = 12, score range 0–12; Cronbach’s alpha = .78)	5.28 / 3.11		
Pearson Test of English General & DIALANG English (<i>k</i> = 54, score range 7–49; Cronbach’s alpha = .85)		30.02 / 8.42	
Pearson Test of English General & DIALANG English (<i>k</i> = 55, score range 10–53; Cronbach’s alpha = .91)			33.59 / 10.05

* *Note.* For statistical analyses, interval scale Rasch scores were used instead of raw scores

APPENDIX C

Descriptive Statistics for the Dependent Measures

Descriptive Statistics for the Cognitive Measures

			4 th grade		8 th grade		Gymnasium	
	min	max	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Backwards digit span in L1	0	14	5.06	1.62	6.57	1.84	7.41	1.86
Backwards digit span in FL	0	8	3.98	1.39	5.61	1.51	5.99	1.39
Rapid recognition of words in L1	0	14	10.32	3.47	13.07	0.88	13.31	0.69
Rapid recognition of words in FL	0	8*	4.69	2.19	11.21	1.33	11.74	0.59
		12**/****						
Rapid word list reading in L1	.42*	1.60**	1.03	0.20	1.36	0.20	1.50	0.19
(words per second; observed min	.73*	1.98**						
& max)	.90***	2.19***						
Rapid word list reading in FL	.37*	1.75*	1.03	0.23	1.51	0.28	1.76	0.29

(words per second; observed min	.95**	2.50**						
& max)	1.12***	2.69***						
Rapidly alternating stimulus in L1	.42*	1.44*	0.88	0.20	0.68	0.16	0.59	0.13
(seconds per item; observed min	.42**	1.32**						
& max)	.32***	.96***						
Rapidly alternating stimulus in FL	.60*	3.40*	1.21	0.41	0.86	0.25	0.68	0.15
(seconds per item; observed min	.32**	1.94**						
& max)	.38***	1.16***						
Pseudo-word reading in L1	0	10	8.76	1.53	9.27	1.15	9.62	0.75
Pseudo-word reading in FL (0–2	0	20	N/A	N/A	14.12	2.76	15.54	2.64
points per item / word)								
Pseudo-word spelling in L1	0	12	9.03	2.00	10.38	1.41	11.22	1.08
Pseudo-word repetition in L1	0	10	6.01	1.46	6.66	1.63	6.99	1.20
Pseudo-word repetition in FL (0–	0	20	10.99	1.74	11.72	1.67	12.45	2.01
2 points per item / word)								
Phoneme deletion in L1 (0–2	0	24	21.30	2.55	22.23	2.19	22.88	1.54

points per item / word)								
Phoneme deletion in FL (0–2	0	12*	6.17	2.61	9.17	2.99	11.70	2.82
points per item / word)		18** / ***						
Common unit in L1	0	10	6.04	2.45	6.50	2.12	6.99	2.33
Common unit in FL	0	10	N/A	N/A	4.66	1.73	5.09	1.73

Note. The minimum and maximum values are the lowest/highest possible score (and the number of items in the test) unless indicated otherwise; *for the 4th graders; **for the 8th graders; ***for the gymnasium students.

Descriptive Statistics for the Linguistic Measures

			4 th Grade		8 th Grade		Gymnasium	
	min	max	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Reading in L1 (Rasch measure)*	−3.31	4.41 (4 th grade)	2.50	1.49	0.70	1.48	1.76	1.30
	−4.09	4.66 (8 th & gymn)						
Writing in L1 (Rasch measure)*	−12.98	4.36 (4 th grade)	−4.54	4.43	−0.87	2.54	4.48	2.88
	−10.25	12.24 (8 th & gymn)						
Writing in FL (Rasch measure)*	−16.92	11.02 (4 th grade)	−1.12	7.17	0.00	3.21	0.01	2.25
	−8.77	7.42 (8 th grade)						

	-6.45	6.19 (gymn)						
Dictation in FL (% correct) (each group had a different test)	0	100 (4 th grade)	71.95	14.21	83.44	16.53	75.10	17.04
	0	100 (8 th grade)						
	0	100 (gymn)						
L1 Vocabulary	0	75	66.35	5.06	71.65	3.40	72.75	3.71
FL Vocabulary	0	60 (4 th grade)	41.60	11.31	64.10	15.90	90.51	19.38
	0	90 (8 th grade)						
	0	120 (gymn)						
Segmentation in L1: time (in minutes)*	0.5	8.0	2.28	1.10	2.52	0.79	1.98	0.67
Segmentation in L1: accuracy (number of errors)*	0	32	5.10	4.49	2.74	4.69	1.07	1.79
Segmentation in FL: time (in minutes)*	0.5	5.0	2.18	0.84	2.08	0.67	2.22	0.66

Segmentation in FL: accuracy (number of errors)*	0	74	17.14	11.97	10.07	7.52	7.93	5.39
Detecting spelling errors: speed (number of words in 3½ minutes)	0	100	N/A	N/A	55.49	20.55	81.16	17.36
Detecting spelling errors: accuracy (number of errors found)	0	100	N/A	N/A	50.55	20.68	77.72	18.07

* The minimum and maximum scores for these measures are actual observed scores; for the other measures, the minimum and maximum score represent possible minimum/maximum scores.

*Descriptive Statistics for the FL Motivation Factors**

			4 th Grade		8 th Grade		Gymnasium	
	min	max	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Anxiety	1	5	2.20	1.08	1.81	0.89	1.89	0.84
English FL Self-concept	1	5	3.48	0.97	3.41	1.06	3.26	1.04
Instrumentality	1	5	4.22	0.69	3.58	0.85	3.73	0.82
Intrinsic Interest	1	5	4.30	0.81	3.48	0.88	3.65	0.81
Motivational Intensity	1	5	3.73	0.70	3.27	0.79	3.18	0.77
Parental Encouragement	1	5	4.33	0.83	3.39	0.86	3.40	0.81
Self-regulation	1	5	3.61	0.88	3.20	0.82	3.13	0.73

*4th graders' motivation questionnaire had somewhat fewer items than the 8th graders' and gymnasium students' questionnaire.

APPENDIX D

Multivariate and Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their L1 Finnish Proficiency

MANOVA Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their L1 Finnish Proficiency

	Weak EFL Readers <i>N</i>	Strong EFL Readers <i>N</i>	Wilks's Lambda	<i>F</i>	<i>p</i>
4 th grade	22	39	.518	10.238	<.001
8 th grade	23	26	.432	7.693	<.001
Gymnasium	23	25	.501	5.688	<.001

Note. Between-groups degrees of freedom = 1 in all cases.

Univariate Tests of the Differences Between Weak and Strong EFL Readers with Respect to Their L1 Finnish Proficiency

4th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	87.27	6.397	1.491	60	.227	.33
	strong	89.40	6.617				

Segmentation	weak	2.250	1.4205	.333	60	.566	.15
(speed)	strong	2.077	.922				
Segmentation	weak	8.45	6.442	16.239	60	<.001	1.07
(accuracy)	strong	3.62	2.926				
Writing	weak	-7.820	4.045	24.570	60	<.001	1.32
	strong	-2.463	4.057				
Reading	weak	1.704	1.294	26.740	60	<.001	1.38
	strong	3.367	1.155				

Note. Between-groups degrees of freedom = 1 in all cases.

8th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	93.45	4.493	17.225	48	<.001	1.21
	strong	97.54	2.123				
Segmentation (speed)	weak	2.783	.795	4.124	48	.048	.59
	strong	2.327	.774				
Segmentation (accuracy)	weak	4.35	6.506	2.870	48	.097	.50
	strong	1.96	2.877				
Detecting spelling errors (speed)	weak	50.61	19.045	5.719	48	.021	.70
	strong	64.92	22.425				
Detecting spelling	weak	42.83	19.913	9.323	48	.004	.89

errors (accuracy)	strong	61.35	22.253				
Writing	weak	-1.784	3.098	2.529	48	.118	.47
	strong	-.503	2.541				
Reading	weak	-.196	1.055	25.724	48	<.001	1.48
	strong	1.627	1.408				

Note. Between-groups degrees of freedom = 1 in all cases.

Gymnasium

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	95.59	4.484	15.899	47	<.001	1.15
	strong	99.25	.949				
Segmentation (speed)	weak	2.437	.662	28.597	47	<.001	1.55
	strong	1.620	.362				
Segmentation (accuracy)	weak	1.43	1.409	6.792	47	.012	.75
	strong	.52	1.005				
Detecting spelling errors (speed)	weak	73.48	15.415	11.120	47	.002	.96
	strong	87.92	14.589				
Detecting spelling errors (accuracy)	weak	69.43	14.292	12.353	47	.001	1.02
	strong	85.00	16.220				
Writing	weak	3.707	2.461	.535	47	.468	.21
	strong	4.308	3.153				
Reading	weak	1.307	1.288	11.425	47	.001	.98

strong	2.493	1.141
--------	-------	-------

Note. Between-groups degrees of freedom = 1 in all cases.

APPENDIX E

Multivariate and Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their English Skills Other than Reading

MANOVA Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their English Skills Other than Reading

	Weak EFL Readers <i>N</i>	Strong EFL Readers <i>N</i>	Wilks's Lambda	<i>F</i>	<i>p</i>
4 th grade	24	38	.386	17.852	<.001
8 th grade	20	26	.124	56.565	<.001
Gymnasium	20	26	.103	69.297	<.001

Note. Between-groups degrees of freedom = 1 in all cases.

Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their English Skills Other than Reading

4th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	34.08	10.056	77.177	61	<.001	2.33
	strong	51.87	6.568				
Segmentation (speed)	weak	2.083	.817	.108	61	.743	.09
	strong	2.013	.818				
Segmentation (accuracy)	weak	28.75	14.624	50.454	61	<.001	1.88
	strong	8.97	7.209				
Writing	weak	-5.566	6.677	27.656	61	<.001	1.39
	strong	3.429	6.487				
Dictation	weak	63.06	14.906	64.449	61	<.001	2.13
	strong	87.47	9.076				

Note. Between-groups degrees of freedom = 1 in all cases.

8th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	44.00	10.498	167.922	45	<.001	3.86
	strong	80.85	8.780				
Segmentation (speed)	weak	2.325	.950	5.710	45	.021	.71
	strong	1.827	.423				
Segmentation	weak	18.95	7.163	83.651	45	<.001	2.72

(accuracy)	strong	4.19	3.578				
Writing	weak	-4.482	1.995	152.041	45	<.001	3.92
	strong	3.082	1.882				
Dictation	weak	56.03	15.556	173.292	45	<.001	3.67
	strong	95.56	4.543				

Note. Between-groups degrees of freedom = 1 in all cases.

Gymnasium

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Vocabulary	weak	64.00	11.947	283.398	45	<.001	5.01
	strong	110.38	6.524				
Segmentation (speed)	weak	2.350	.366	9.695	45	.003	.93
	strong	1.808	.708				
Segmentation (accuracy)	weak	14.15	5.833	80.714	45	<.001	2.67
	strong	2.65	2.591				
Writing	weak	-2.944	1.594	120.990	45	<.001	3.27
	strong	2.912	1.926				
Dictation	weak	48.83	12.643	298.107	45	<.001	5.14
	strong	94.52	4.224				

Note. Between-groups degrees of freedom = 1 in all cases.

APPENDIX F

Multivariate and Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to the Cognitive Factors

MANOVA Analyses of the Differences Between Weak and Strong EFL Readers with Respect to the Cognitive Factors

	Weak EFL readers <i>N</i>	Strong EFL readers <i>N</i>	Wilks' Lambda	<i>F</i>	<i>p</i>
4 th grade	21	32	.655	4.958	.001
8 th grade	20	26	.531	9.057	<.001
Gymnasium	23	18	.534	7.860	<.001

Note. Between-Groups degrees of freedom = 1 in all cases.

Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their Cognitive Skills

4th Grade

Variable (Cognitive Factor)	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Fluent reading 1	weak	-.303	.993	17.303	52	<.001	1.19
	strong	.697	.754				
Fluent reading 2	weak	-.056	.824	6.924	52	.011	.75

	strong	.604	.934				
General phonology	weak	-.383	.717	4.663	52	.036	.62
	strong	.078	.785				
Working memory	weak	-.032	1.094	.153	52	.698	.11
	strong	.071	.824				
Specific Finnish phonology	weak	.051	.725	.585	52	.448	.22
	strong	.215	.791				

Note. Between-Groups degrees of freedom = 1 in all cases.

8th Grade

Variable (Cognitive Factor)	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Fluent reading	weak	-.956	1.053	19.473	45	<.001	1.34
	strong	.243	.721				
English phonology	weak	-1.242	1.223	26.296	45	<.001	1.56
	strong	.201	.785				
Working memory	weak	-1.095	1.157	32.885	45	<.001	1.74
	strong	.378	.543				
Finnish phonology	weak	-.648	1.553	3.652	45	.063	0.58
	strong	.001	.680				

Note. Between-Groups degrees of freedom = 1 in all cases.

Gymnasium

Variable (Cognitive Factor)	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Fluent reading	weak	.099	.783	9.754	40	.003	.98
	strong	.819	.660				
English phonology	weak	-.117	.718	23.669	40	<.001	1.53
	strong	.894	.579				
Working memory	weak	-.251	.668	27.998	40	<.001	1.69
	strong	.859	.639				
Finnish phonology	weak	.042	.528	5.546	40	.024	.75
	strong	.469	.626				

Note. Between-Groups degrees of freedom = 1 in all cases.

APPENDIX G

Multivariate and Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to the Motivational Factors

MANOVA Analyses of the Differences Between Weak and Strong EFL Readers with Respect to the Motivational Factors

	Weak EFL readers <i>N</i>	Strong EFL readers <i>N</i>	Wilks's Lambda	<i>F</i>	<i>p</i>
4 th grade	13	31	.703	2.172	.060
8 th grade	16	21	.442	5.230	.001
Gymnasium	21	21	.256	14.134	<.001

Note. Between-Groups degrees of freedom = 1 in all cases.

Univariate Analyses of the Differences Between Weak and Strong EFL Readers with Respect to Their Motivational Factors

4th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Anxiety	weak	.324	1.407	2.450	43	.125	.47

	strong	-.209	.837				
English FL self-concept	weak	-.466	1.204	11.470	43	.001	1.03
	strong	.653	.839				
Instrumental motivation	weak	-.075	1.534	.303	43	.585	.17
	strong	.124	.855				
Intrinsic interest	weak	-.100	1.503	1.396	43	.244	.36
	strong	.321	.850				
Motivational intensity	weak	.073	1.596	.154	43	.696	.12
	strong	.217	.840				
Parental encouragement	weak	.246	1.258	2.339	43	.134	.46
	strong	-.279	.940				
Self-regulation	weak	.077	1.432	.476	43	.391	.21
	strong	-.150	.940				

8th Grade

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Anxiety	weak	.029	.879	1.076	36	.307	.35
	strong	-.267	.840				
English FL self-concept	weak	-.702	1.064	33.310	36	<.001	1.97
	strong	.868	.571				
Instrumental motivation	weak	-.801	1.099	8.606	36	.006	1.00

	strong	.158	.890				
Intrinsic interest	weak	-.755	1.069	5.719	36	.022	.82
	strong	.071	1.018				
Motivational intensity	weak	-.360	1.056	1.231	36	.275	.38
	strong	.034	1.083				
Parental encouragement	weak	-.290	1.214	.046	36	.831	.07
	strong	-.214	.917				
Self-regulation	weak	-.318	1.104	.150	36	.701	.13
	strong	-.164	1.255				

Gymnasium

Variable	Groups	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	<i>d</i>
Anxiety	weak	.839	1.368	15.339	41	<.001	1.24
	strong	-.466	.680				
English FL self-concept	weak	-1.165	.762	95.496	41	<.001	3.09
	strong	.922	.614				
Instrumental motivation	weak	-.194	1.048	.839	41	.385	.29
	strong	.101	1.035				
Intrinsic interest	weak	-.472	.641	26.193	41	<.001	1.61
	strong	.670	.797				
Motivational intensity	weak	-.636	.731	8.867	41	.005	.94

	strong	.225	1.104				
Parental encouragement	weak	-.244	1.153				
	strong	-.050	.913	.366	41	.549	.19
Self-regulation	weak	-.290	.662				
	strong	.101	.997	2.237	41	.143	.47

APPENDIX H

Factor Loadings for Exploratory Factor Analysis (Principal Axis Factoring) with Promax Rotation and Kaiser Normalization of the Cognitive Measures

4th Grade

Measure	Fluent Reading 1	Fluent Reading 2	General (L1 & FL) Phonology	Working Memory	Specific Finnish Phonology
Rapidly presented words in English	.825	.451	.143	.315	.297
Rapidly presented words in Finnish	.736	.436	.248	.255	.356
Rapid alternating stimulus naming in English	.536	.498	.252	.164	.221
Word list reading in English	.766	.840	-.282	.307	.298
Word list reading in Finnish	.744	.775	.085	.386	.547
Rapid alternating stimulus naming in Finnish	.300	.566	.080	.152	.154
Common phonological unit in Finnish	.340	.449	.271	.258	.432

Pseudo-word repetition in English	.156	.321	.311	.045	.171
Pseudo-word spelling in Finnish	.205	.210	.532	.218	.405
Pseudo-word repetition in Finnish	.179	.078	.509	.208	.266
Phoneme deletion in English	.333	.268	.474	.139	.158
Backwards digit span in Finnish	.180	.144	.099	.758	.192
Backwards digit span in English	.371	.323	.199	.594	.282
Pseudo-word reading in Finnish	.194	.152	.245	.183	.713
Phoneme deletion in Finnish	.202	.269	.274	.108	.360

Note. Factor loadings > .400 are in boldface.

8th Grade and Gymnasium

Measure	Fluent	Phonological Skills		Phonological Skills
	Reading	Mainly in English	Working Memory	Mainly in Finnish
Word list reading in English	.862	.527	.412	.240
Word list reading in Finnish	.827	.428	.362	.294

Rapid alternating stimulus naming in				
English	.746	.645	.564	.265
Rapid alternating stimulus naming in				
Finnish	.737	.410	.364	.319
Phoneme deletion in English	.434	.716	.435	.351
Pseudo-word reading in English	.411	.588	.477	.430
Pseudo-word repetition in English	.240	.522	.249	.177
Rapidly presented words in English	.389	.432	.423	.161
Pseudo-word repetition in Finnish	.237	.419	.161	.329
Pseudo-word spelling in Finnish	.277	.355	.147	.323
Rapidly presented words in Finnish	.283	.329	.291	.117
Backwards digit span in English	.296	.368	.716	.214
Backwards digit span in Finnish	.271	.281	.570	.250
Common phonological unit in Finnish	.306	.329	.493	.326
Phoneme deletion in Finnish	.268	.372	.367	.695
Pseudo-word reading in Finnish	.250	.261	.216	.613

Common phonological unit in English	.132	.244	.229	.318
-------------------------------------	------	------	------	------

Note. Factor loadings > .400 are in boldface.