PREDICTING READING ACHIEVEMENT IN A TRANSPARENT ORTHOGRAPHY: RUSSIAN CHILDREN LEARN TO READ

A Dissertation
Submitted
to the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education

By
Katerina Petchko
January, 2009

ABSTRACT

This study investigated the cognitive, linguistic, and reading skills of 79 Russian-speaking first and second graders to determine the strongest concurrent predictors of reading achievement. The children were administered a battery of 15 tests from which nine objective, interval-scale measures were derived: phonological awareness, verbal short-term memory, decoding accuracy, listening comprehension, reading comprehension, nonverbal ability (IQ), vocabulary, decoding rate, and rapid naming. In a series of multiple regression analyses, phonological awareness accounted for a small amount of unique variance in both decoding accuracy and decoding rate whereas rapid naming was a unique predictor of decoding rate only. Neither verbal short-term memory nor IQ accounted for any variance in decoding. For reading comprehension, IQ and linguistic comprehension contributed a substantial amount of variance to the prediction of achievement whereas decoding rate did not. However, in a series of direct discriminant function analyses, reliable differences emerged between good and poor decoders on reading comprehension, indicating that decoding *is* relevant to reading comprehension in young Russian readers. The good and poor readers also differed reliably on phonological awareness, vocabulary, IQ, and listening comprehension. Taken together, the findings indicate that the cognitive operations that drive the development of reading in the Russian language share many similarities with those in other languages including English. However, they also indicate some important characteristics of Russian reading that make it distinct from reading in other languages. These

characteristics concern the relative weight of cognitive abilities in determining reading achievement, the manifestations of reading disability, and the reading strategies that Russian children appear to adopt for word recognition and may reflect the effect of the largely transparent Russian orthography. These findings converge with those obtained in other languages and suggest that reading in all alphabetic languages may be characterized by universal as well as language-specific aspects. This study also investigated the structure of phonological awareness in the Russian language, revealing, through Rasch analysis, that this ability is best conceptualized as a unidimensional ability that can be measured by a variety of tasks.

ACKNOWLEDGMENTS

The completion of this research required the practical, intellectual, and philosophical contribution from so many people. It is with deep gratitude and appreciation that I acknowledge them here.

To Dr. David Beglar, my adviser: Thank you for taking me under your wing four years ago and for believing in this project from its inception. Thank you for the continuous guidance, commitment, and insightful advice at every stage, and for the careful and perceptive reading of at least a dozen drafts. I could not wish for a better adviser. You have taught me to think like a scientist, to read like a scholar, and to be precise with words; you have helped me grow professionally and shown me what academic excellence really means. Above all, you have helped me find a definitive answer to the question that, at some point, faces many doctoral students: How much do I really want this degree? For that, I am eternally grateful.

To Dr. Kenneth Schaefer, TESOL Program Director: Thank you for your support throughout my studies at TUJ. Without it, completing the doctoral program would have been so much more difficult. It has been an honor and a privilege to study under you and I am so grateful for the exceptional kindness and infinite patience you have shown me.

To Dr. Mark Sawyer: It was in your Psychology class that I first became interested in reading research and dyslexia. You have helped me shape my interest in the topic, define my research goals, and reconcile my research interests with the constraints of the program and I thank you for that. Thank you also for the valuable

suggestions on my manuscript and for the feedback at my proposal defense and, later, my dissertation defense. I have benefited tremendously from your criticism.

A huge round of thanks to Dr. Steven Ross and Dr. Donna Tatsuki for graciously sharing their knowledge and insights with me and for providing helpful suggestions in refining my work. I am also grateful to Dr. JD Brown for helping me appreciate the elegance of statistical reasoning, the power of statistical analysis, and the limitations of statistical models.

To Professor Tomoko Nemoto: Thank you for your practical help during my six years at TUJ and, especially, for your advice at a crucial stage in my research. You have gone far beyond the call of duty to make my experience at TUJ less stressful and more meaningful and I thank you from the bottom of my heart.

To Professor Donna Amoroso: Thank you for supporting me during the past four years, for being so flexible, patient, and understanding, and for pushing me to complete this research. Your support and encouragement have helped me to focus, to keep going, and to overcome the many hurdles I encountered as a doctoral student with a full-time job, a family, and ailing parents. I feel that I have been truly blessed by having you as a colleague and a friend.

To Dr. Steven Connolly, my best cohort buddy: Your wit, humor, and healthy doses of cynicism have helped me keep my eyes on the ball and my feet on the ground. Thank you for your friendship!

To Arnold Fisher, Anna Burova, and Ronald Magallanes: Thank you for the encouragement and emotional support and for being such wonderful and faithful

friends. It is in no small measure due to your caring, loyalty, and compassion that I survived these past six years. (The alcohol helped too!)

Warm thanks to the principals and teachers of the participating schools and the children who participated in this study. Thank you for giving up valuable class time to allow me to collect data. Without you, there would be nothing to report. I am also very grateful to Dr. Ludmila Gorevaya and Ms. Alla Barannick for helping me with the testing.

To my dear parents, Dr. Valeryi G. Petukhov and Dr. Vera G. Rodionova: Папа, спасибо за твою неоценимую практическую помощь во время сбора и обработки данных, спасибо за советы и за множество интересных идей. Без твоей помощи я никогда не смогла бы даже начать, не то что закончить. Мама, спасибо за помощь в поиске школ, в получении разрешения на тестирование, и в поиске ассистентов. Без твоей организации не о чем было бы писать. Но больше всего, спасибо вам, дорогие родители, за моральную поддержку, за заботу, и за бесконечную любовь. Я вам очень благодарна!

Finally, to Pavel and Takaji: You are the ones to whom I owe the most. Thank you for sharing with me the frustrations, challenges, and doubts throughout this project. Takaji, спасибо за твою бесконечную заботу и за то, что ты так много взял на себя, чтобы я могла учиться. Pasha, спасибо тебе за техническую помощь, за понимание и безотказность, и за то, что столько раз выручал меня в трудной ситуации. Но больше всего спасибо за то, что ты такой замечательный сын! I dedicate this work to you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The Background of the Issue

Reading is a complex cognitive activity that draws upon the reader's ability to process and integrate phonological, orthographic, and semantic information (Adams, 1990; Ehri, 2002; Knight & Hynd, 2002). Fluent reading relies on two processes, word recognition and meaning comprehension, which must operate in unison (Ehri, 1995, 2002). These two processes can be clearly differentiated, and they appear to draw upon somewhat different cognitive abilities (de Jong & van der Leij, 2002). Yet, they are also closely related as there is a strong and positive connection between the efficiency of word recognition and the accuracy of reading comprehension (Stanovich, Cunningham, & Freeman, 1984; Tunmer, Herriman, & Nesdale, 1988). Whether this connection is causal (e.g., de Jong & van der Leij) or merely correlational remains to be seen; what is clear at this time is that the ability to read and comprehend depends upon the rapid and automatic recognition of single words. This process is seen as central in many theoretical accounts of fluent reading development (Ehri, 1995; Snowling, 2000).

From a cognitive, behavioral, and neurobiological view, the process of word recognition exhibits striking similarities across alphabetic languages: It involves similar cognitive and metalinguistic mechanisms (Frost, 1998; Goswami, 2000; Perfetti, 2003; Seidenberg, 1992); it depends on the use of a limited number of

orthography-to-phonology translation strategies (Ehri, 2002); and it relies on a common network of primarily left-sided language areas of the brain (Paulesu, McCrory, Fazio, Memoncello, Brunswick, Cappa, et al., 2000). Moreover, a body of evidence is accumulating suggesting that the *acquisition* of reading skills in different languages is based on similar cognitive principles and involves similar mechanisms, although there are important language-specific influences as well (Frost, 1998; Pugh, Sandak, Frost, Moore, & Mencl, 2005). The similarities exist because word recognition in alphabetic languages is essentially phonological (Ehri, 1995; Ziegler & Goswami, 2006), and it depends on the mastery of an "arbitrary code of conventions" (Turner, 1997, p. 16) used by a given culture to represent speech through visual symbols. To become literate, therefore, children must learn to map these visual symbols (or graphemes) onto units of sound (or phonemes), a process called *phonological recoding*[1] (Share, 1995).

---

[1] In the reading literature, several terms have been used to describe the process of converting printed letter strings into speech-based information including *phonological coding, phonological recoding*, and *decoding*. These terms are often used interchangeably but may also be interpreted somewhat differently depending on the context. In the present study, the term *phonological recoding* is used to denote the process by which speech-based information is derived from visual symbols and it is reserved to the discussion of the theoretical accounts of reading and reading development, where it is contrasted with *visual-orthographic processing*. The term *decoding,* on the other hand, is used as an umbrella term to refer to a group of processes by which readers achieve word recognition including the process of *grapheme-phoneme decoding*, or transforming letters into sounds and blending the sounds to form words, and the process of *reading by analogy*, or generating new-word pronunciations from stored pronunciations of orthographically similar words.

Yet, at all three levels—the cognitive, the behavioral, and the neurobiological—there appear to be important differences in the way readers of different languages process print. These differences manifest themselves in the patterns of brain activation during reading (Fiez, 2000; Paulesu et al., 2000), in the choice of specific reading strategies (Ziegler & Goswami, 2006; Ziegler, Perry, Jacobs, & Braun, 2001), as well as in the relative weight accorded to a particular mechanism or process in a given language (see Frost & Katz, 1992, for a review). Whether these differences are a matter of degree or of kind is still being debated; what appears unequivocal is that they are a result of the core characteristics of the spoken and written language in use.

Languages vary along several important dimensions including phonological, morphological, and syntactic. More important, languages vary in orthographic depth, or the consistency with which they use letter symbols to represent individual phonemes. In some orthographies, one letter can have several pronunciations, whereas in others, individual letters or letter clusters are always pronounced in the same way. Similarly, in some orthographies, one phoneme or phoneme cluster can be spelled in many different ways, whereas in others, a phoneme is always represented by the same grapheme. These differences become crucial when children begin learning to read.

Three major conclusions can be drawn from the findings of cross-linguistic research on reading development in alphabetic languages. First, there appears to be a strong connection between the degree of orthographic consistency of a given

language and the ease of reading acquisition in that language: Children learning to read a consistent orthography acquire the grapheme-phoneme conversion rules needed for fluent reading much more quickly than do children learning to read an inconsistent orthography (Aro & Wimmer, 2003; Seymour, Aro, & Erskine, 2003). For example, in English, one of the most orthographically inconsistent alphabetic languages, the rate of reading development is more than twice as slow as in consistent orthographies such as German, Greek, or Italian (Seymour et al., 2003).

Second, recent cross-linguistic comparisons indicate that the very nature of word reading may be different in consistent versus inconsistent orthographies. For example, recent experiments conducted by Ziegler and his colleagues (Goswami, Ziegler, Dalton, & Schneider, 2003; Ziegler et al., 2001) in English and German as well as studies on the use of specific strategies in reading English and other languages (e.g., Goswami, 1988; Holopainen, Ahonen, & Lyytinen, 2002) suggest that the significant differences in reading performance observed across languages are a result of the differences in the reading strategies employed by readers in different orthographies. To explain these cross-linguistic differences, Ziegler and Goswami (2005) proposed a framework that they called *a psycholinguistic grain-size theory*. According to this theory, children learning to read a consistent orthography rely almost exclusively on grapheme-phoneme decoding, whereas children learning to read an inconsistent orthography have to supplement grapheme-phoneme decoding with rhyme and whole-word reading strategies in order to cope with orthographic inconsistency. Ziegler and Goswami have argued

that the reading system develops differentially in different languages in response to the demands of the particular orthography.

Third, language-specific properties appear to determine not only the course of reading development but also the manifestations of difficulty in readers with specific reading disability, or *dyslexia*. It has been shown that orthographically inconsistent languages impose a much greater difficulty on dyslexic readers than do orthographically consistent languages (see Goulandris, 2003b, for a review). Moreover, the patterns of difficulty exhibited by dyslexic readers in inconsistent orthographies appear to be qualitatively different from the patterns exhibited by such readers in consistent orthographies. Numerous studies have shown that in inconsistent orthographies such as English, dyslexia typically presents as a reading accuracy impairment, with accuracy being "at the heart of the matter for most dyslexia assessments" (Turner, 1997, p. 109; see also Snowling, 2000, for a summary of dyslexia research in English), whereas in consistent orthographies, such as German or Greek, dyslexia is characterized primarily by slow, dysfluent, and effortful reading (Landerl, 2001; Nikolopoulos, Goulandris, & Snowling, 2003; Wimmer, 1996). Other linguistic features such as morphology and syntax may also influence the types of problems encountered by dyslexic readers and the extent to which these problems may interfere with reading performance (e.g., Abu-Rabia, 2004; Gyarmathy, 2004; Leong, 1999; Share & Leikin, 2004; Smythe & Everatt, 2000).

Given the cross-linguistic variation in the development of skilled reading as well as the varying patterns of difficulty associated with reading development in different languages, it is perhaps unsurprising that the relative importance of cognitive abilities that have been implicated in reading development varies greatly across languages. For example, phonological awareness, or the sensitivity for speech sounds in spoken words, has been shown to be a strong and reliable predictor of reading attainment in English across school grades (see Goswami, 2000, 2002, Snowling, 2000; Ziegler & Goswami, 2005, for reviews); however, its predictive relation to reading in other languages is weaker than that of other abilities (Holopainen, Ahonen, & Lyytinen, 2001; Wimmer, Mayringer, & Landerl, 2000) or is limited to a short period (de Jong & van der Leij, 2002; Frost, Madsbjerg, Niedersoe, Olofsson, & Sorensen, 2005).

Similarly, the specific tasks that predict reading disability or that discriminate between good and poor readers most reliably vary across orthographies (Everatt, Smythe, Adams, & Ocampo, 2000; Everatt, Smythe, Ocampo, & Gyarmathy, 2004), and tasks designed for one language may not be sufficiently sensitive or otherwise appropriate for other languages (Nikolopoulos et al., 2003; see Smythe et al., 2004, for a review). These findings have prompted some researchers to conclude that it is difficult to generalize reading assessment procedures from one linguistic context to another.

The ability to recognize words fluently, however, is only half of the "reading equation"; the other half is the ability to comprehend the meaning of print.

According to one influential view of reading, the Simple View (Gough & Tunmer, 1986; Hoover & Gough, 1990), reading comprehension can be described by the relationship between word decoding and linguistic comprehension.[2] The predictive validity of this relationship has been demonstrated in a number of languages including English (Joshi & Aaron), Dutch (de Jong & van der Leij, 2002), and French (Megherbi, Seigneuric, & Ehrlich, 2006). However, there is growing evidence that the linguistic features of a language may affect the relative weight of decoding versus linguistic comprehension in determining reading comprehension. For example, in English, reading comprehension in beginning readers has been found to be more closely associated with decoding than with linguistic comprehension (e.g., Hoover & Gough; Shankweiler et al., 1999), whereas in French (Megherbi et al., 2006), the opposite relationship has been observed.

Orthographic complexity also affects the strength of the relationship between reading speed and reading comprehension. In a cross-linguistic comparison, Oney and Goldman (1984) found that Turkish children showed stronger and more stable correlations between the speed of word reading and the accuracy of passage comprehension than did American children of similar age and ability. The authors argued that for languages with a higher degree of letter-sound

---

[2] Linguistic comprehension is sometimes referred to in the literature as *oral language comprehension.* Both listening comprehension and vocabulary tests have been used, sometimes separately and sometimes as a composite measure, to assess linguistic comprehension.

correspondence, greater automaticity with decoding is associated with better comprehension. This may be because in such languages, decoding accuracy is attained quite rapidly after the onset of reading instruction (Seymour et al., 2003) and reading speed remains the only useful criterion in differentiating good from poor readers.

The Research Problem

The above discussion suggests that despite the commonalities in the basic psychological operations applied in oral reading in all languages (Besner & Smith, 1992; Seidenberg, 1992), there are important orthography-specific aspects of reading and reading acquisition that make it difficult to generalize models of skilled as well as impaired reading across languages. This is an important point because most current theoretical accounts of reading and reading acquisition have been developed for English. Yet it appears that learning to read English is not only disproportionately harder than learning to read more consistent orthographies like German, Greek, or Italian but also, possibly, qualitatively different (Ziegler et al., 2003). Many manifestations of cognitive deficits associated with reading disability also appear to be language-specific (see Goulandris, 2003a, 2003b, for a review), and no universal profile of dyslexia has so far been described.

Under these circumstances, language-specific models of reading and reading disability become all the more important, both for the advancement of reading theory and for practical considerations, including the development of

appropriate reading tests and dyslexia screening instruments. However, although research on reading in non-English languages has intensified considerably in the past decade, there are still many languages that have not been studied systematically. Russian is one such language. This is regrettable because Russian is the mother tongue for some 270 million people (Crystal, 1992) and the language of instruction or the preferred language of communication for millions more. How do Russian children learn to read? Is this process similar to that in English? What cognitive abilities are most closely associated with the development of both word decoding and reading comprehension in Russian and what measures best discriminate between good and poor readers? This study attempts to answer these and related questions.

Purposes of the Study

The first purpose of this study is to explore concurrent predictors of reading in Russian-speaking beginning readers. Specifically, I wish to explore the relative influence of phonological awareness, verbal short-term memory, and rapid naming in the prediction of decoding accuracy and decoding speed in Russian in order to identify the strongest independent predictors and to determine to what extent measures developed for other languages can also be applied to Russian. A second purpose is to examine the structure of phonological awareness in the Russian language in order to determine whether linguistically and cognitively diverse tasks form a single variable. A third purpose is to explore the relative influence of

nonverbal ability, linguistic comprehension (i.e., listening comprehension and vocabulary), and decoding skill on reading comprehension in order to test the predictive power of the Simple View of reading in Russian. The final purpose is to examine the cognitive and linguistic skills of good and poor readers of Russian in order to determine their relative strengths and weaknesses and to attempt to identify, among the poor readers, children who may be suffering from specific reading disability, or dyslexia.

Significance of the Study

This study is the first attempt to explore the process of normal reading development in the Russian language systematically in order to identify its similarities to, and differences from, the process of reading development in other languages, particularly in English, and place Russian reading acquisition within the framework of existing theories of reading and cross-linguistic reading research. As such, this study adds to the scholarly literature on reading development in different orthographies. Furthermore, by benchmarking the course of normal reading development in the Russian language, this study contributes to the development of culture- and language-appropriate tools for evaluating reading achievement and for identifying children in Russian schools who are at risk for reading failure so that appropriate intervention and remediation strategies can be developed. This study also attempts to clarify the main sources of difference between good and poor readers and to examine the cognitive and reading profile of children who may be

classified as dyslexic. The particular strength of this study is in the use of truly objective measures of cognitive, linguistic, and reading ability derived from the Rasch measurement model.

The Audience for the Study

The audience for this study includes researchers who are interested in exploring the theoretical aspects of cross-linguistic reading acquisition and especially the effect of orthography on reading development; elementary school teachers who teach literacy skills; school psychologists who test children entering elementary school; test developers who are interested in constructing culture-appropriate tests to measure various cognitive and linguistic abilities in Russian children; writers of Russian language and Reading textbooks; and school administrators and educational policy-makers who are interested in improving literacy instruction as well as testing in Russian schools.

Delimitations

The focus of this study is on the cognitive and behavioral aspects of reading and reading development. More specifically, in investigating word reading, I focus on the phonological, processing speed, and memory aspects of decoding and do not investigate motor or sensory (e.g., visual/auditory) dimensions. In investigating reading comprehension, I focus on vocabulary knowledge, listening comprehension, and decoding automaticity and do not examine other possibly important factors

such as metacognition or the role of context. Further, I focus on the concurrent predictors of reading acquisition, which may differ from those that predict reading achievement longitudinally.

The investigation of the behavioral aspects of reading is limited to cultural-linguistic factors, primarily orthography, and does not consider such potentially important environmental factors as the learning context, the social setting, and the educational authority. The focus of the study is on unselected native Russian-speaking elementary school children who are learning to read in their native language.

CHAPTER 2

REVIEW OF THE LITERATURE

Reading in Alphabetic Languages

The Phonological Nature of Reading

The goal of reading is to extract meaning from printed text. To be successful, the reader must first retrieve the meaning of each individual word he or she encounters, put the meanings of these words together, and interpret the whole phrase within the context of the sentence, paragraph, and the overall understanding of the text (Adams, 1990). At the very bottom of this process lies fluent and automatic word recognition.

Skilled readers read at a rate of 250 words per minute, or four words per second (Aaron & Joshi, 1992), and they appear to recognize the majority of words that they read holistically, by sight. Yet, eye movement studies (see summaries in Aaron & Joshi, 1992, and Adams, 1990) indicate that skilled reading involves the relatively complete processing of the individual letters of words as even very slight misspellings create apparent disruptions in normal readers' eye movements. So how can readers read as fast as 250 words per minute and still process each individual letter? To answer this question, it is important to examine what happens in normal reading.

A widely held view (e.g., Adams, 1990) is that skilled word recognition involves three components: orthographic processing, phonological processing, and

lexical processing. The thousands of words that mature readers have in their mental lexicons are represented in terms of their orthographic, phonological, and lexical properties. To recognize these words, readers must identify their orthographic patterns, compute their phonological codes, and retrieve their meaning from long-term memory. This process is shown schematically in Figure 1 (modified from Adams, 1990). The three processors shown in the figure are not independent--they work together on the same operation, guiding and facilitating the work of each other. Together, these processors form the reading system.

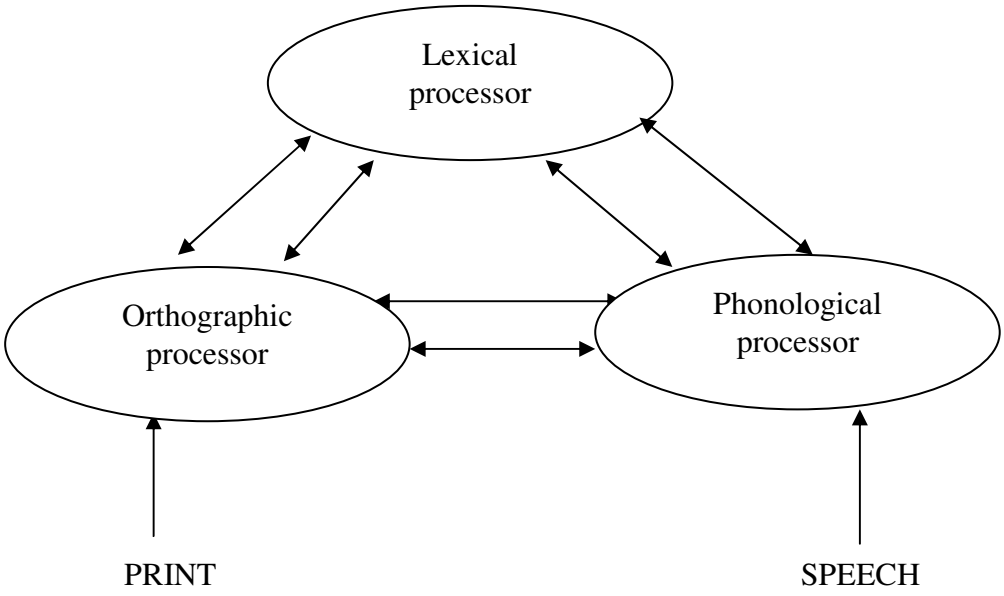


Figure 1. A framework of lexical processing in alphabetic languages.

In alphabetic orthographies, this system assumes the existence of two pathways to meaning—a direct orthographic pathway and a phonologically

mediated pathway. Just how prevalent one of these pathways is relative to the other is an open question. According to the traditional, dual-route, conception of reading (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Paap, Noel, & Johansen, 1992; Seidenberg, 1992), word frequency is the primary factor that determines whether a word is recognized on a visual-orthographic or phonologically mediated basis: Familiar words are processed visually whereas unfamiliar words require the assembly of a phonological code based on the reader's knowledge of sound-spelling correspondences. Implicit in this conception is the superiority of the visual-orthographic strategy, which enables readers to recognize words rapidly and automatically by sight.

Although the exact mechanism of sight word reading is not fully understood, some theorists (Ehri, 1995; 2002) have argued that readers learn sight words by forming connections between graphemes in the spellings and phonemes in the pronunciations of individual words, which link written words to their pronunciations and meanings in memory. When readers see words whose orthographic and phonological information they have stored in their mental lexicons, the shape of these words activates their pronunciation and meaning and enables their recognition as holistic patterns, without the need for letter-to-sound transformation. Evidence to support sight word reading comes from experimental observations showing that children can read familiar words such as *book* much faster than they can read simple nonwords such as *mig* (e.g., Ehri & Wilce, 1983).

Other theorists working in the area of reading research have argued, however, that the phonological structure of a word is the main link to the word's meaning and that meaning retrieval without phonology is impossible. For example, Frost (1998) describes a phonological model of reading in which the default operation of the cognitive system in word recognition is the prelexical assembly of phonology. According to Frost, access to a phonological representation always precedes meaning retrieval, and orthography affects semantic comprehension through phonological processing, not directly. Theoretically, this argument rests on an observation that all writing systems are phonological in nature and their primary purpose is to represent spoken language (Perfetti, 2003). In this view, the phonological code of a word is computed by the cognitive system automatically upon the visual presentation of a word.

This position has received considerable empirical support. Numerous experiments have shown that the use of phonology in word identification is a fundamental, and perhaps an automatic component of skilled reading: It is observed in reading aloud, in silent reading, and even in reading tasks in which it hinders performance (Lukatela & Turvey, 1994; Perfetti, Bell, & Delaney, 1988; Tzelgov, Henik, Sneg, & Baruch, 1996; Van Orden, Johnston, & Hale, 1988; see also Frost, 1998, for a review of the experimental paradigms used to investigate the use of phonology in reading). For example, the work of Van Orden and his colleagues using semantic categorization tasks suggests that word recognition proceeds "from spelling to sound to meaning" (Van Orden, Johnston, & Hale, p. 371) and that

meaning activation is always phonological. In their experiments, participants were asked to categorize words and nonwords, some of which were homophonic to category exemplars (e.g., *hare* as a part of the human body) while others looked similar in spelling (e.g., *harp*). The researchers found that the homophones produced far more false positive categorizations than did the spelling controls. They interpreted this result as evidence for the obligatory involvement of phonology in meaning activation.

Automatic phonological activation in word recognition has been observed in languages other than English, for example, in Hebrew (Frost & Kampf, 1993), Arabic (Bentin & Ibrahim, 1996), and Chinese (Perfetti & Zhang, 1991, 1995). These findings have led some researchers (e.g., Frost, 1998) to conclude that phonological activation in reading for meaning occurs in all languages, irrespective of their linguistic properties, and that it is a universal feature of reading. Furthermore, recent research has demonstrated the crucial importance of phonological recoding to the *early stages* of reading acquisition in virtually all orthographies in which reading development has been studied (see below).

Learning to Read

Several stage-based models have been proposed to describe reading development in alphabetic languages (e.g., Ehri, 1995; Frith, 1985). In the main, these models assume a developmental shift from the use of phonological recoding in the early stages to the use of direct orthographic processing once reading

proficiency has been attained. For example, in Ehri's model, reading acquisition begins as a non-alphabetic process in which pre-readers establish connections between certain visual cues (such as the golden arches behind the "McDonalds" sign) and whole words. As children learn how letters in written words represent sounds in their pronunciations, reading acquisition changes into an alphabetic process, involving establishing connections, through phonological recoding, between letters and sounds (Ehri & Wilce, 1985). At first, these connections link only the most salient letters to sounds; later, as learners acquire greater alphabetic knowledge—typically as a result of print experience—these connections become more fine-grained. Spelling patterns that recur regularly in different words become consolidated into multi-letter patterns symbolizing phonological blends. These patterns accumulate in memory and form the reader's generalized knowledge of the spelling system, which facilitates the learning of new words by sight.

The problem with stage-based models, however, is that they do not fit well with experimental evidence demonstrating that beginning and skilled readers alike can and do use both direct orthographic and phonological recoding strategies in word recognition (Martin, Claydon, Morton, Binns, & Pratt, 2003) and that even beginning readers can read simple words by sight (Ehri & Wilce, 1983). Share (1995) proposed an alternative model in which phonological recoding is not stage-based but item-based—whether it is employed depends on the reader's familiarity with the encountered word. High-frequency words, which have been encountered previously, can be recognized visually virtually from the onset of reading

acquisition. In contrast, low-frequency words that are unfamiliar to the reader will require the use of phonological recoding regardless of the reader's skill level. Because young readers are more likely to encounter many unfamiliar words, they will rely on phonological recoding to a much greater extent than will skilled readers.

In this "phonology by familiarity" (Share, 1995, p. 97) model, phonological processing acts as a self-teaching device, enabling the child to acquire the detailed orthographic representations necessary for fast and accurate word recognition. In any given text, a majority of words is likely to be high-frequency and familiar to the reader and will therefore be recognized visually. However, a small number of low-frequency items will be unfamiliar and will require phonological recoding, thereby providing opportunities for the reader to acquire the specific orthographic representations of these words. Thus, within this model, phonological recoding represents "the *sine qua non* of reading acquisition" (Share, p. 156) because it affects the speed and quality with which orthographic representations are acquired. In this model, the contribution of visual/orthographic factors to the acquisition of fluent word recognition is secondary, and it is not equivalent to that of phonological recoding, because "only phonology offers a functional self-teaching mechanism" (p. 169). In a later study, Share (1999) tested this model in Hebrew and found support for the self-teaching role of phonology in the acquisition of word recognition skills. It is unclear, however, to what extent this model can be

generalized to other orthographies, particularly those with more predictable sound-to-spelling correspondences.

A third view of reading acquisition sees this process as a continuum in which readers become increasingly more efficient in phonological recoding. Frost (1998), for example, has argued that both beginning and skilled readers rely on phonological recoding in reading and that the only difference between proficient and less proficient readers is in the efficiency of this process. As Frost pointed out, this course of development is consistent with cognitive models of general human learning, in which extensive practice with a cognitive skill results in greater efficiency rather than in a shift to a different routine.

Although interesting theoretically, the question of whether reading acquisition can in fact be conceptualized as progressing from a greater reliance on phonological recoding to a greater reliance on orthographic processing as the reader becomes more competent or whether achieving reading fluency is a matter of becoming more efficient at computing phonological codes goes beyond the scope of this study. What is important and what appears to be widely acknowledged in the reading literature is that phonological processing is routinely involved in reading and that achieving reading proficiency depends heavily on the successful operation of the phonological component. More important, previous research has stressed the development of efficient access to phonological representations of words in fluent reading in most languages studied so far (e.g., Goswami, 2000, 2002; see also Goulandris, 2003b, for a review). These findings have prompted some researchers

to suggest that the cognitive principles and operations underlying the development of word recognition in reading are very similar in all orthographies (Besner & Smith, 1992; Frost, 1998; Perfetti, 2003; Seidenberg, 1992). However, there also appear to be important differences in how written languages are processed and represented in the brain (Goswami et al., 2003; Ziegler et al., 2001), which may significantly affect the course of reading development. These differences have been attributed to the differences in linguistic features between languages, particularly to the differences in orthography.

The Effect of Orthography on Reading

Orthography represents the spoken language. As such, it maps systematically onto phonology rather than meaning (Perfetti, 2003). Yet the way different orthographies represent the phonology of the spoken language varies across languages. First, there are differences in the size of the basic speech unit denoted by one written symbol (i.e., a phoneme in alphabetic languages, a character in logographic languages, and a syllable in syllabaries). Second, there are differences in syllabic complexity, which largely reflect the distinction between Romance languages, having a predominance of open syllables and few consonant clusters, and Germanic languages, having many closed syllables and complex consonant clusters (Seymour et al., 2003). Alphabetic writing systems further differ in the degree of reliability of grapheme-phoneme correspondence rules that underpin the consistency of sound-letter mappings. The continuum ranges from

perfectly consistent (also called *shallow*, *regular*, or *transparent*) orthographies in which there is a one-to-one correspondence between letters and sounds (e.g., Finnish) to inconsistent (also called *deep*, *irregular,* or *opaque*) orthographies in which there is a many-to-many correspondence between letters and sounds (e.g., English). In such orthographies, one letter or letter combination may have many different pronunciations, and one sound or sound combination may be spelled in many different ways. In English, for example, about 40 phonemes are transcribed with a mere 26 letters; in contrast, in Russian, one grapheme corresponds (roughly) to a single phoneme (Grigorenko, 2003). This variation has major consequences for the reader.

Many studies have shown that in consistent orthographies such as German or Finnish, even beginning readers rely almost exclusively on assembling word pronunciations on-line, by direct phonological recoding (Frith, Wimmer, & Landerl, 1998; Holopainen et al., 2002; Landerl, Wimmer, & Frith, 1997). In contrast, readers in less consistent orthographies such as English or Hebrew have been shown to rely on direct visual-orthographic strategies (Frost & Bentin, 1992) or to supplement grapheme-phoneme decoding with other strategies such as reading by analogy (Goswami, 1988).

Until recently, the most influential theory explaining this cross-language variation has been the *orthographic depth hypothesis* (Katz & Frost, 1992), which postulates that readers in consistent orthographies, where the mapping between letters and sounds is unambiguous, will rely more on phonological recoding in

reading, whereas readers in inconsistent orthographies will rely more on orthographic processing. The strong form of this theory, which precludes the use of visual-orthographic strategies in perfectly consistent orthographies, has been falsified by evidence showing that phonological processing is routinely involved in reading even the most inconsistent orthographies (e.g., see Katz & Frost, 1992, for a review). However, the weak form, according to which both orthographic and prelexically assembled phonological information can be used for lexical access but the degree to which readers rely on one type of coding or the other depends on orthographic consistency, appears to be plausible on rational grounds and consistent with dual-route models of reading (see the previous section).

More recent evidence, however, suggests that orthography influences not so much the relative contribution of phonology to reading as the very nature of the phonological processes involved. Ziegler et al. (2001) compared the reading of identical words and nonwords in English and German by skilled adult readers. The researchers manipulated word length and body neighborhood (body-*N*), or the number of words sharing the same orthographic rhyme with the target word (e.g., *late, date*, and *fate* are body-*N*s of *hate*). They hypothesized that because of the consistency of German orthography, the German readers would exhibit a stronger word length effect (i.e., reading shorter items faster than longer ones) as a result of letter-by-letter decoding. In contrast, the English readers would show a stronger body-*N* effect (i.e., reading items with small body-*N*s more slowly than items with large body-*N*s) as a result of reading by analogy to compensate for the

inconsistency of the English spelling. This is exactly what they found. Their results suggest that readers in consistent orthographies rely more on small-unit processing, whereas readers in inconsistent orthographies prefer large-unit processing strategies.

These results were replicated in a later study with English and German children (Goswami et al., 2003). The researchers gave the children three lists of nonwords: one containing familiar large-unit orthographic patterns that could be read most successfully by applying a large-unit strategy (e.g., *dake* [*cake, make*]), one containing unfamiliar large-unit orthographic patterns that could be read most successfully by decoding (e.g., *daik*), and one containing both types of pattern. They predicted that in reading the third list, German children would be able to rely exclusively on decoding whereas English children would have to switch between decoding and large-unit processing, thereby incurring a switching cost. This prediction was confirmed.

Results such as these demonstrate that orthographic differences between languages generate qualitatively distinct processing strategies for word recognition. These differences are captured well in the *psycholinguistic grain-size theory* (Ziegler & Goswami, 2005), which stipulates that the size of the psycholinguistic units that develop for reading varies and depends on the linguistic features of a given language (e.g., orthographic consistency). According to this theory, readers in consistent orthographies can be very successful using only a small-unit grapheme-phoneme decoding strategy because of orthographic consistency at the

small grain size level, whereas readers in inconsistent languages have to rely on several strategies for reading: a small-unit strategy to cope with the visual complexity of large orthographic clusters, and a large-unit strategy to cope with the grapheme-phoneme inconsistency.

Recent brain-imaging studies (e.g., Fiez, 2000; Paulesu et al., 2000) provide additional evidence to support the above findings. These studies suggest that while the organization of the brain architecture that supports reading appears to be similar in speakers of different languages, the degree of activation of particular neural networks in reading varies as a function of orthographic consistency. For example, in a comparison of English and Italian readers, Paulesu and his colleagues observed different patterns of brain activation: The Italian subjects showed greater activation in areas associated with phoneme processing, whereas the English subjects showed greater activation in areas associated with word retrieval and naming tasks. These differences were attributed to the variation in processing demands associated with the orthographic consistency of the writing system.

The Effect of Orthography on Learning to Read

In alphabetic languages, the relationship between the letters making up a particular word and the word's meaning is always arbitrary. In contrast, the relationship between letters and sounds is systematic (Frost, 1998; Ziegler & Goswami, 2005). In other words, knowing that a particular word contains certain letters tells the reader nothing about its meaning; it does, however, enable the

reader to access the pronunciation of the word. Learning how to read in an alphabetic language, then, involves learning how the letters in printed words map onto the sounds in spoken words, or establishing connections between written language (orthography) and spoken language (phonology). Mastery of these connections enables children to read the thousands of words already present in their spoken lexicon.

Recent empirical research suggests that the ease with which children master the connections between phonemes and graphemes differs across languages and depends on the particular orthography to be learned. Seymour et al. (2003) reported the results of a project conducted by scientists in 14 European countries in order to compare the rates of acquisition of word recognition and decoding skill in children after one year of reading instruction. Scientists participating in the project developed a matched set of items that included simple familiar words to measure word recognition and simple pronounceable nonwords to measure decoding ability. They found that both the accuracy and speed of reading varied considerably among orthographies as a function of orthographic consistency and, in the case of nonword reading, of syllabic complexity. English turned out to be the hardest language to learn, with an average accuracy rate of 34% for word reading and 29% for nonword reading. The average accuracy for the other languages ranged from 98% (Finnish) to 71% (Danish) for word reading and from 95% (Finnish) to 53% (Danish) for nonword reading.

The researchers also modeled, for the individual languages, a normal range of reading development defined as plus or minus 1.75 *SD* from the mean. In most cases, the normal range for the inconsistent orthographies overlapped the disability range for the consistent orthographies, extending in some cases into the non-reader zone. The researchers interpreted the greatest delay in reading acquisition observed in English as indicative of a combined effect of its syllabic complexity and orthographic inconsistency on reading development. They speculated that reading acquisition in a consistent orthography may be based on a single decoding ability whereas in an inconsistent orthography, it may require both decoding and visual-orthographic processing. This division of processing resources, according to Seymour and his colleagues, results in slower learning.

Similar results have been obtained in other studies involving cross-language comparisons (e.g., Aro & Wimmer, 2003; Gholamain & Geva, 1999; Oney & Goldman, 1984; Wimmer & Goswami, 1994; Wimmer & Hummer, 1990). These studies have consistently shown that in consistent orthographies, children have little problem translating letter strings—even unfamiliar ones—into pronunciations and they quickly achieve reading fluency, whereas in inconsistent orthographies, beginning readers often experience difficulty with fast and accurate word decoding, and on average, take much longer to acquire orthographic-phonological relations. Comparisons of error rates in nonword reading are particularly informative in this respect: In English after one year of reading instruction, the mean error rate in nonword reading tasks is around 40% (Jorm, Share, MacLean, & Matthews, 1984;

Juel, Griffith, & Gough, 1986); in contrast, in German, the error rate for comparable samples is around 10% (Frith et al., 1998; Wimmer & Hummer, 1990).

There are also qualitative differences in the process of reading acquisition in consistent versus inconsistent orthographies, reflecting differences in the nature of phonological recoding, as has been shown by Frith and her colleagues (1998). In a longitudinal study of English and German children, the researchers examined word and nonword reading in 7- to 12-year olds and found that the acquisition of phonological recoding in English was not only much slower than in German but also that it followed a qualitatively different path: The orthographic inconsistency of English did not permit the online assembly of syllables and necessitated the use of complex and error-prone strategies in word recognition, whereas the orthographic consistency of German enabled efficient phonological recoding of syllables online.

Additional evidence for the differential development of reading skills in different orthographies comes from cross-language studies of children exhibiting a significant discrepancy between their actual reading performance and that predicted based on their general intellectual ability as reflected in their IQ test scores. Such children are believed to be suffering from *dyslexia* (see pp. 34-36). It has been argued (e.g., Goulandris, 2003a; Nikolopoulos et al., 2003; Zeigler & Goswami, 2005) that the severity of reading impairment in dyslexic children depends on the orthography a child has to master. In fact, even though dyslexics are almost always impaired compared to their controls (e.g., Goulandris, 2003b), those reading a

consistent orthography typically outperform those reading an inconsistent orthography on a matched set of items (Landerl et al., 1997). For example, in Landerl et al.'s study, the German dyslexics read difficult three-syllable words more accurately than the English dyslexics read simple one-syllable words.

There are also qualitative differences in the way that dyslexia manifests itself in different orthographies. In consistent orthographies, dyslexics typically exhibit a marked speed deficit on reading tasks (Lyytinen, Aro, & Holopainen, 2004; Reitsma, 2004; Wimmer, 1996), whereas in inconsistent orthographies, dyslexia can be diagnosed on the basis of accuracy problems alone (Snowling, 2000; Turner, 1997), although speed problems are also characteristic features of this impairment. Furthermore, an accurate assessment of cognitive deficits underlying dyslexia may require the use of more sensitive measures in consistent orthographies than in inconsistent orthographies (Everatt et al., 2004; Nikolopoulos et al., 2003), as tasks designed for inconsistent orthographies such as English may be too easy for readers of more consistent orthographies. This finding in particular has important implications for cross-linguistic reading research, as it suggests that cognitive measures designed for assessing reading performance in one linguistic environment (e.g., English) may not be appropriate in another (e.g., Russian).

In summary, the following conclusions can be drawn from the above review. First, although there are universal aspects of reading that concern the basic processing mechanisms, reading strategies that readers adopt for word recognition appear to be language-specific rather than universal. Second, the process of reading

acquisition in consistent orthographies appears to differ both quantitatively and qualitatively from the same process in inconsistent orthographies. These differences are evident in the rate of acquisition of grapheme-phoneme conversion rules as well as in the strategies children develop for reading. In particular, the decoding process, which is seen as central in theoretical accounts of reading acquisition, develops more slowly and less effectively in English than in other alphabetic languages and has to be supplemented by other strategies for fluent reading. Third, reading problems appear to manifest themselves differently in consistent versus inconsistent orthographies and detecting these problems calls for the use of language-specific tests.

These conclusions warn against the uncritical adoption and use of reading models and measures across different orthographies. If the reading system develops differently in response to the orthographic features of the particular language, then the reading models developed for English (e.g., dual-route models) may not be appropriate for languages like Russian, in which grapheme-phoneme correspondences are relatively straightforward and consistent. It is also possible that the specific set of cognitive and linguistic abilities that correlate with reading, the structure of these abilities, and their relative importance to reading development will vary as a function of orthographic consistency. This possibility is of particular concern given the enormous impact theoretical models of reading in English have had on reading research in other languages.

Russian Orthography

Russian is an Indo-European language that belongs to the Eastern Slavic subgroup. It is written in the Cyrillic alphabet in which some letters are written and pronounced approximately as in English (e.g., *m, k*), some are written as in English but are pronounced differently (e.g., *o, p*), and some are written and pronounced differently (e.g., *ё, ц*). The Russian alphabet consists of 33 letters: 21 for consonants, 10 for vowels, and two letters without sound—a soft sign and a hard sign.

The phonological system of Russian is similar to that of English in a number of key dimensions. As in English, the smallest sound unit is the phoneme. Russian has a syllable structure similar to that of English and it allows consonant clusters in both syllable onsets (e.g., *СТРОИЛ* /stroil/ "[he] was building") and codas (e.g., *ГОСТЬ* /gos't'/ "guest"). Russian words have one strong stress, which consists of volume and length, although the stress pattern of multi-syllabic words is much more predictable than in English (Hamilton, 1980). As in English, Russian vowels are reduced in unstressed positions, so the same vowel letters may be pronounced differently in different positions.

There are also important differences. Most Russian consonants come in "hard/soft" pairs. For example, the first sound of the word *СОН* (/son/ "dream") is hard, while the first sound of the word *СЕЛ* (/sel/ "[he] sat down") is soft, although both sounds are represented by the same letter. Each Russian vowel is represented by two letters—one for the vowel after a hard consonant, the other for the vowel

after a soft consonant. For example, /mat'/ ("mother") is written as *МАТЬ*, while /myat'/ ("to crumple") is written as *МЯТЬ*. Hamilton (1980) maintains that in Russian phonetics, "consonants are king, and vowels are slave" (p. 33): The quality of the vowel is determined by the quality of the consonants on either side of it. In Russian spelling, however, the situation is just the opposite: The vowel letters provide information about the softness or hardness of the preceding consonant, and so in Russian spelling, vowels are king. Hamilton argues that this allows maximum efficiency as just 10 vowel letters can be used to represent all hard and soft consonants.

How transparent is Russian orthography? According to Hamilton (1980), this can be established by answering two questions: Would knowing the spelling of a Russian word guarantee knowledge of its pronunciation? and Would knowing the pronunciation of a Russian word guarantee knowledge of its spelling? Hamilton's answer to the first question is "almost always," with a few exceptions. His answer to the second question is "occasionally," with the majority of words having some sounds that could be represented by alternative letters.

The basic principle of Russian orthography is morphological in that each morpheme is spelled in the same way despite its different pronunciations in different word forms. In Russian, the true identity of a vowel is the quality that it would show under stress, so to know how to spell an unstressed vowel, one must find a morphologically related word where this vowel is in a stressed position. For example, the word *ГОРОДОК* (/gorodok/ "small town") has main stress on the last

syllable, so the first two vowels are reduced and pronounced [ə] and [a], respectively. To know how to spell the first vowel, one needs to find a related word where this vowel would be in a stressed position, for example, the word *ГОРОД* /gorod/ "city"; to know how to spell the second vowel, another related word must be found, with the main stress on the second "o," for example, the word *ИНОГОРОДНИЙ* (/inogorodnyi/ "out-of-town"). The problem, however, is that there are many words whose spelling cannot be verified in this way. These words are called exceptions and their spelling must be memorized.

Russian spelling is also tied to semantics, which helps readers distinguish between homophones. For example, the second, unstressed, vowel in the words *УМАЛЯТЬ* (/umalyat'/ "to belittle") and *УМОЛЯТЬ* (/umalyat'/ "to beg") is pronounced the same, as [a] and both words sound identical. To determine how to spell these words, one must consider what they mean in the particular context and then find a morphologically related word with the vowel in question in a stressed position.

There are also many deviations from the morphological principle of spelling, which are associated with tradition. For example, some adjective or verb endings are traditionally spelled in a particular way; the spelling of many foreign words borrowed into Russian often deviates from their pronunciation. However, as Hamilton (1980) argues, many irregular spelling patterns occur in predictable and familiar environments and can be learned quickly. English has many more archaic spelling patterns than Russian.

In summary, Russian orthography is relatively consistent, with most phonemes represented by only one letter and most letters representing only one phoneme, with a few exceptions. However, Russian orthography also has elements of inconsistency, particularly in spelling, where some letters can have more than one phonemic realization. To put it differently, Russian orthography can be described as relatively consistent for reading and relatively inconsistent for spelling, although Russian spelling is much more systematic than English. For more information about Russian orthography, see Hamilton (1980) and Grigorenko (2003).

Specific Reading Disability: Dyslexia

Why Do Some Children Fail to Learn to Read?

To learn to read, children need to appreciate that printed forms on a page correspond to spoken language. To do that, they need to set up a system of mappings between letters and sounds, which is based on the alphabetic principle. This alphabetic principle is crucially important because it provides the child with a self-teaching device that enables the decoding of unfamiliar words (Share, 1995). A child who has grasped the alphabetic principle can abstract the correspondence between the letters and the sounds in familiar words and use this knowledge to read unfamiliar words (Snowling, 2000).

Yet, in all alphabetic languages[3] studied so far, a small group of children (e.g., 2-11%, depending on the language, Smythe & Everatt, 2000) lag significantly behind their peers in reading development—or fail to learn to read altogether—despite normal or above normal intelligence, adequate reading instruction, and no general learning disabilities, sensory deficits, emotional problems, or socioeconomic disadvantage. These children are said to be suffering from dyslexia, or specific reading disability. The specific criteria for diagnosing dyslexia vary across countries, populations, and studies, but typically a discrepancy of about two standard deviations between a child's actual reading ability and that predicted based on their IQ (Fawcett, 2002) must be demonstrated before that child can be classified as dyslexic. Reading problems in dyslexic children are manifested in extreme difficulty with the acquisition of very basic reading skills such as single word reading and letter-sound decoding.

According to the highly influential *phonological representations hypothesis* of dyslexia (e.g., Goswami, 2000; Snowling, 2000), dyslexic children's difficulty with reading acquisition stems from their phonological deficits, or "weak phonological [re]coding" (Vellutino, Fletcher, Snowling, & Scanlon, 2004, p. 12), which prevents them from establishing fine-grained links between graphemes in

---

[3] Readers interested in learning about dyslexia in non-alphabetic languages, particularly Japanese, are directed to Hirose and Hatta (1988), Wydell (2003), Wydell and Butterworth (1999), Wydell and Kondo (2003), Yamada (2004), and Yamada and Banks (1994).

written words and phonemes in spoken words. Their representations of word pronunciations remain coarse-grained, perhaps at the level of whole words, which makes it impossible for such children to develop a self-teaching device for reading unfamiliar words. As Snowling (2000) has put it, "a dyslexic child who learns to pronounce *pat* as 'pat' will, when faced with *cap*, have to learn this anew" (p. 89).

Dyslexia and Nonword Reading

The most direct test of phonological recoding skill is a test of nonword reading in which a child is presented with pronounceable nonsense words and asked to read them as fast and as accurately as possible. Because nonwords are visually unfamiliar, the only way to read them is by using phonological recoding. This test, then, becomes the benchmark measure of phonological recoding, giving an indication of the integrity of a child's phonological skills.

Most studies conducted with English-speaking dyslexics have reported nonword deficits in dyslexic readers relative to both reading-age (RA) and chronological-age (CA) controls. For example, in a comprehensive review of studies of nonword reading in English, Rack, Snowling, and Olson (1992) reported differences between dyslexic and RA-matched controls ranging from 9 to 43 percentage points, with a median of 19 percentage points. They further showed that studies that had failed to find any differences in nonword reading between dyslexic and normal children were methodologically deficient and did not control for differences in IQ between the two groups, the difficulty of the nonwords used, or

the similarity of the test items to real words. In a meta-analytic study using the same database, Van Ijzendoorn and Bus (1994) examined these factors systematically and confirmed that the main reason for the failure to obtain differences on nonword reading between dyslexic and normal readers was the inadequacy of the matching procedure in terms of age, intelligence, and word recognition ability. They also confirmed the presence of a nonword reading deficit in dyslexia, concluding that it would take more than 400 studies with null results "to bring the combined probability...[of no such deficit] down to insignificance" (p. 274).

Nonword reading deficits in dyslexia have been reported in languages other than English. For example, Greek first graders with dyslexia read 93% of nonwords correctly compared with 97% for CA controls (Porpodas, 1999); Dutch first graders read an average of 4.8 nonwords correctly, compared with 6.7 nonwords for the non-dyslexic group—a significant difference (de Jong & van der Leij; 2003); French dyslexics read 75% of nonwords correctly compared with 90% for CA controls (Sprenger-Charolles, Colé, Lacert, & Serniclaes, 2000). In a longitudinal English-German comparison using identical nonwords, Frith et al. (1998) found that English children made more errors than German children at all times at which they were tested and even when word-recognition ability was equated between the groups. Landerl et al. (1997, see also Landerl, 2003, for a summary of English-German comparisons) tested 10- and 12-year-old English and German dyslexics on nonword reading and found that the English group were much less accurate than

the German group, producing, for example, 14 different pronunciations for a few three-syllable nonwords.

It may be misleading, however, to base estimates of decoding ability in consistent orthographies on nonword reading accuracy alone because the straightforward grapheme-phoneme correspondences in such orthographies make it possible even for children with weak phonological skills to achieve high levels of decoding accuracy. For example, Wimmer et al. (2000) reported that German children with a phonological deficit scored about 90% correct on a test of easy nonwords and about 76% correct on a test of difficult nonwords—in both cases, only about 4% lower than the no-deficit group. In another German study (Wimmer, 1996), ten-year-old dyslexics read simple and complex nonwords at 91% and 84% accuracy rates, respectively. In Dutch, de Jong and van der Leij (2003) found that although normal first graders were significantly more accurate on nonword reading than dyslexic children, there was a considerable overlap (about 50%) in accuracy scores between the groups.

In contrast, nonword reading speed appears to be a strong and reliable differentiator between dyslexic and normal readers in both consistent and inconsistent orthographies. For example, in de Jong and van der Leij's (2003) study described above, the nonword reading speed scores of the dyslexic and normal readers did not overlap at all and differed by more than 4.5 standard deviations. In Wimmer's study (1996), the dyslexic group showed marked speed impairments compared to RA controls both on simple nonwords derived from frequent content

words (e.g., *put, ema*) and on more difficult nonwords derived from Japanese (e.g., *matoru, kerata*). In Ziegler et al.'s (2003) study, both English and German dyslexics showed impairments in nonword reading speed with respect to both CA and RA controls; however, neither group was impaired on accuracy relative to RA controls. More important, the speed deficit was of similar size in both languages. The authors concluded that the nonword decoding speed deficit affects dyslexics in both consistent and inconsistent orthographies in similar ways. A very similar pattern was observed in a study by Paulesu and his colleagues (2001), who compared Italian, French, and English dyslexic adults on nonword reading speed. Despite a consistent advantage in absolute terms exhibited by the Italian readers over the English and French groups, when the *z*-scores were compared across the groups, the Italian dyslexics differed from their controls by as much as the English and French groups differed from *their* controls.

Correlates of Word Reading

Phonological Awareness

Phonological awareness can be defined as one's awareness of, and access to, the sound structure of oral language (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). It involves the ability to notice, identify, think about, and manipulate phonological segments within words including syllables (e.g., *tic-tac*), onset-rime units (e.g., *gr-asp*), and phonemes (e.g., /*s-i-t*/). It is typically demonstrated by such tasks as detecting a non-rhyming word in a string of three or four words, tapping

out the number of syllables or phonemes in a word, isolating initial or final sounds in a word, or deleting a phoneme from a word (Adams, 1990). Thus, depending on the task, the term *phonological awareness* has been used in developmental literature to refer to syllabic awareness, rime (or rhyme)[4] awareness, and phoneme awareness. Phonological awareness is a metacognitive ability as it requires conscious reflection (Snowling, 2000).

A clear developmental progression in phonological skills has been demonstrated in most alphabetic languages studied so far (see Ziegler & Goswami, 2005, for a review). The development proceeds from a shallow sensitivity to large units (e.g., syllables and onset-rime units) to a deep sensitivity to small units (i.e., phonemes). According to this view, which is exemplified in the work of Goswami (1999, 2000, 2004), it is the syllable, and not the phoneme, that is the most basic unit of speech processing across all languages; syllable-based reading is thus seen as the most basic.

The close connection between children's phonological awareness and the ease with which they acquire basic word recognition skills has been well documented (see Adams, 1990; Snowling, 2000; Wagner & Torgesen, 1987, for reviews). This relationship has been demonstrated across orthographies and in bilinguals (see Comeau, Cormier, Grandmaison, & Lacroix, 1999, for French; de

---

[4] In the literature on reading, rime refers to the linguistic unit of the syllable containing the vowel and the succeeding consonant or consonants, for example, *-at* in the word *cat*. Rhyme, in contrast, refers to the shared phonetic similarity of the end sounds of at least two words, for example, *coat* and *goat*. In practice, these terms often overlap.

Jong & van der Leij, 2003, for Dutch; Frost et al., 2005, for Danish; Gottardo, 2002, for Spanish-English bilinguals; Mannai & Everatt, 2005, for Arabic; McBride-Chang & Kail, 2002, for Chinese; Oktay & Aktan, 2002, for Turkish; Porpodas, 1999, for Greek; Share, 2003, for Hebrew). Some theorists (e.g., Elbro, 1996) have suggested that phonological awareness is the single strongest predictor of reading achievement, and many now believe that the relationship between phonological awareness and reading is causal rather than correlational (Share, 1995; Veluttino et al., 2004; Wagner & Torgesen, 1987; Ziegler & Goswami, 2005).

Several lines of evidence support the critical role of phonological awareness in reading acquisition. The first one comes from studies demonstrating strong predictive relations between phonological awareness and early reading skills. Particularly informative in this respect are longitudinal studies. Bradley and Bryant (1983) followed the progress of some 400 children for four years starting from when the children had not yet started to read. They found a strong relationship between the children's phonological awareness at age four and their reading skills at age eight. Muter and Snowling (1998) reported a study of 34 children tracked for five years. The children completed tests of rhyme awareness (rhyme detection and discrimination) and phoneme awareness (phoneme deletion) when they were four, five, six, and nine years old. Hierarchical multiple regression analyses showed that phoneme awareness was a significant predictor of individual variation in reading accuracy, both concurrently and longitudinally. In another study, Nation and

Snowling (2004) followed the development of reading skills in 72 eight- to 13-year olds and found that concurrently, measures of phonological awareness and nonword decoding accounted for an impressive 72% of the variance in word recognition and that longitudinally, they accounted for 10% of the variance in word recognition.

The second line of evidence comes from studies of poor readers, or readers with dyslexia. Numerous studies have shown that children who demonstrate impairments in the phonological domain—particularly on tasks requiring the manipulation of phonemes—also have difficulty learning to read (Elbro, Borstrom, & Petersen, 1998; Landerl et al., 1997; Marshall, Snowling, & Bailey, 2001; see also Goulandris, 2003b; Snowling, 2000, for reviews). The strongest designs have compared dyslexics with chronological and reading-age controls in order to highlight developmental deficits that are the cause rather than the consequence of dyslexics' reading problems (Abu-Rabia, Share, & Mansour, 2003; Bruck, 1992; Pennington, Van Orden, Smith, Green, & Haith, 1990; Swan & Goswami, 1997). In a recent comprehensive review of some 40 years of research on dyslexia, Vellutino et al. (2004) have concluded that deficiencies in phonological awareness are causally related to problems with word recognition. This conclusion is further supported by case studies of developmental dyslexia (Campbell & Butterworth, 1985; Ramus et al., 2003; Snowling, Hulme, & Goulandris, 1994).

Perhaps the strongest support for the role of phonological awareness in reading comes from training studies in which phonological awareness has been

manipulated experimentally. These studies have consistently shown positive effects of phonological training not only on the development of phonological awareness but also on reading achievement. For example, Lundberg, Frost, and Petersen (1988) trained a group of Danish preschoolers to attend to, and produce, rhymes, syllables, and phonemes before these children started to read. They found significant effects of the training on the children's performance on rhyming tasks and an even greater effect on the children's performance on phoneme tasks. Moreover, significant facilitating effects were observed on the children's progress in reading in Grades 1 and 2. Phoneme segmentation and phoneme isolation training significantly improved reading skills in a sample of Norwegian children (Lie, 1991), while phoneme awareness instruction combined with letter name and letter sound teaching had a beneficial effect on the development of early reading skills in English-speaking children (Ball & Blachman, 1991). In a meta-analysis of nearly two dozen training studies with a randomized or matched design conducted in the U.S., Bus and Ijzendoorn (1999) found that phonological awareness training explained about 12% of the variance in reading skill. Positive effects of phonological training on reading development have also been found in highly orthographically consistent languages such as German (Schneider, Kuspert, Roth, & Vise, 1997).

Many of the above studies have demonstrated that the effect of phonological awareness training is specific to reading and does not extend to other linguistic skills such as vocabulary (Lundberg et al., 1988), or to mathematics

(Bradley & Bryant, 1983). This suggests that phonological awareness is not simply a predictor of general academic achievement but rather a reading-related skill.

The Effect of Language Characteristics on Phonological Awareness

With the emergence of cross-linguistic studies of reading, it has become increasingly evident that the relationship between phonological awareness and reading development is not the same in different orthographies. For example, Holopainen et al. (2001) found that in Finnish, phonological awareness tasks failed to predict delays in reading development in poor readers. De Jong and van der Leij (2002) followed a group of 140 Dutch children from Grade 1 to Grade 3 in order to examine the development of their phonological abilities and found that the effect of phonological awareness on reading was limited to the first grade, after which phonological awareness did not provide an extra impetus for the development of word recognition. The design of their study was particularly strong as they took into account the autoregressive effects of prior reading. Similarly, in Danish, phonological awareness was the strongest predictor of reading only in the early stages of reading development (at the start of Grade 2); at all other times (Grades 2-9), semantic skills were the strongest predictor (Frost et al., 2005).

Cross-linguistic studies also indicate that the manifestation of phonological awareness deficits depends on the orthography a child has to master and that such deficits are generally much harder to detect in consistent than in inconsistent orthographies. For example, Porpodas (1999) found that the performance of Greek

dyslexics on phonological awareness tasks—although worse than that of controls—was nevertheless highly satisfactory, at 78-88% correct, depending on the task. He attributed this result to the regular nature of the Greek orthography. In Hungarian, another orthographically consistent language, phonological awareness tests failed to distinguish good from poor readers (Everatt et al., 2004). De Jong and van der Leij (2003) followed a group of Dutch dyslexics from kindergarten through Grade 6 and found that by Grade 6, deficits in phonological awareness disappeared in this group. They attributed these results to the regular nature of the Dutch orthography.

The characteristics of the spoken language also appear to exert a significant influence on phonological awareness and its development. For example, it has been shown that children using alphabetic languages such as English perform better on phonological awareness tasks than children in logographic languages such as Chinese (Bialystok, McBride-Chang, & Gigi, 2005). On the other hand, children learning to read a consistent alphabetic orthography have been shown to develop phonological awareness much faster and more easily than children learning to read an inconsistent orthography (see Goswami, 2002, for a review). In an Italian-English comparison, Cossu, Shankweiler, Liberman, Katz, and Tola (1988) found that phoneme awareness was more developed in Italian than in English-speaking children. In a Turkish-English comparison, Durgunoglu and Oney (1999) reported that the Turkish speakers outperformed the English speakers on both phoneme and syllable tasks. The faster development of phonological awareness in consistent languages is often attributed to their simpler phonological and syllabic structure.

The influence of the spoken language on phonological awareness may be even more fundamental as spoken language characteristics appear to affect, through readers' preference for syllable segmentation, the very basic sensitivity to phonological units. For example, according to Treiman (1985, 1986), in English, the most natural way to divide syllables is into onsets (i.e., the sounds preceding the vowel) and rimes (i.e., the vowel and consonant(s) following it): *sp-in, sh-eep*, which makes English children highly sensitive to rime units in learning to read. However, in Korean, the preferred syllabic structure of these words would be *spi-n* and *shee-p* (syllable body plus coda), which makes Korean readers more sensitive to body-coda units than to onset-rime units (Perfetti, 2003). Perfetti has argued that this difference is at the linguistic level, not just the reading level, because it parallels preferences in the division of spoken words. My own preliminary work with adult native speakers of Russian indicates that Russian speakers overwhelmingly prefer body-coda segmentation to onset-rime segmentation. Such linguistic differences become crucial in selecting appropriate tasks to measure phonological awareness.

Phonological Awareness Tasks

Task difficulty. A number of tasks have been used to measure phonological awareness. Adams (1990) identified at least five levels of difficulty of phonological tasks differing from one another by the type of cognitive operation required to complete a task (e.g., detection, blending, segmentation, deletion, or substitution)

or by the size of the unit of sound that is the focus of the task (i.e., whether a task requires the identification or manipulation of syllables, intrasyllabic units such as rhymes, or phonemes).

At the lowest level is the knowledge of nursery rhymes, which according to Adams (1990) requires nothing more than an ear for the sounds of words. At the next level are rhyme and alliteration tasks, which require the ability to focus on certain sounds. These are followed by phoneme blending and syllable splitting, tasks that require an understanding that words can be divided into smaller units. At the next level is phoneme segmentation, which requires the ability to analyze words into a series of phonemes. At the highest level of difficulty are phoneme manipulation tasks, which require not only the ability to subdivide a word into phonemes but also the ability to manipulate them by addition, deletion, or transposition.

Several researchers investigated the issue of task difficulty empirically. For example, Yopp (1988) examined the relative difficulty of ten phonological tasks including rhyming, auditory discrimination, phoneme blending, word matching, phoneme segmentation, sound isolation, phoneme deletion, and phoneme counting. Because the tasks varied in the number of items, she converted the means to a percent correct for each test. She found that the easiest tasks were rhyming, auditory discrimination, and phoneme blending, and the hardest—phoneme segmentation and phoneme deletion.

Stahl and Murray (1994) also examined the levels of difficulty of common phonological tasks, but obtained different results. They administered four tests varying in the type of operation required (i.e., phoneme blending, segmentation, isolation, and deletion) to 113 kindergarten and first-grade children and found that after equating for linguistic complexity, phoneme isolation was the easiest task, followed by blending, deletion, and segmentation. In terms of linguistic complexity, the easiest items were those that required analyzing onsets and rimes, followed by those that required analyzing vowels and codas, followed by those that required analyzing cluster codas, and finally, cluster onsets.

One explanation for the contradictory results obtained by Yopp (1988) and Stahl and Murray (1994) is that the type of cognitive operation required to complete a task and the level of linguistic complexity are not the only factors that influence task difficulty. Other task characteristics such as the position of the target linguistic unit within the word, the position of the target word within a string of words, or the nature of the target word (i.e., whether it is a real or a nonsense word) may also affect task difficulty. A recent study using item response theory (IRT) found empirical confirmation of this hypothesis (see below).

Schatschneider, Francis, Foorman, Fletcher, and Mehta (1999) used an IRT model to show that a task's difficulty depended on many factors including the size of the linguistic unit to be manipulated, the position of the target linguistic unit within the word, the type of operation to be performed, and the nature of the target word. For example, the difficulty of deleting phonemes from a word depended on

whether the phoneme to be removed occurred in the middle of the word (i.e., mean difficulty = .89), or at the beginning or end of the word (mean difficulty = -.54); the difficulty of blending phonemes depended on whether the sounds to be blended formed a real word or a nonsense word; the difficulty of categorizing sounds depended on the position of the target word within a string of words, with items with the target word in the final position being consistently easier than other items. These results show that performance on an item can be substantially influenced not only by task requirements but also by the characteristics of the target item and the nature of the target word. Schatschneider et al. also found that a task's usefulness in estimating a child's phonological awareness depended on the child's age, most likely reflecting the child's developmental maturity. For example, tasks requiring a child to blend onsets and rimes were useful in estimating phonological awareness of younger children, but did not fare well in estimating phonological awareness of older average-ability children.

The predictive power of phonological awareness tasks. Another important question is whether various phonological awareness tasks are equally important in predicting reading development. Traditionally research in this area has compared the usefulness of tasks measuring the processing of small units (i.e., phonemes) versus the usefulness of tasks measuring the processing of large units (i.e., rhymes). Results for the usefulness of rhymes have been mixed. For example, in English, rhymes have been found to be the more powerful predictor of reading achievement

in a number of studies (for a comprehensive review, see Goswami, 1999), especially in younger children, prompting some researchers (e.g., Goswami & Bryant, 1990) to claim that there are causal connections between pre-school rhyme awareness and later progress in reading.

However, other researchers did not find rhyme awareness useful in predicting reading ability, either concurrently or longitudinally (Hulme, Hatcher, Nation, Brown, Adams, & Stuart, 2002; Muter & Snowling, 1998; Nation & Hulme, 1997). For example, in Muter and Snowling's study, rhyme awareness failed to predict reading ability in a group of preschool and elementary school children either in the short or in the long term. In Nation and Hulme's study, onset-rime segmentation ability failed to predict any shared or unique variance in reading ability in elementary school children. An interesting finding in that study was that the children found onset-rime segmentation tasks very difficult, performing them at about 55% correct. Nation and Hulme argued that this finding seriously undermines Goswami and Bryant's claim that onset-rime awareness develops early and naturally.

Results for orthographically transparent languages have been more consistent, typically showing weak or non-existent relations between onset-rime awareness and reading development. For example, Hoien, Lundberg, Stanovich, and Bjaalid (1995) investigated predictors of reading in six- and eight-year-old Norwegian preschoolers and first-grade students. Although the rhyme measure did make a statistically significant contribution to predicting reading ability in the

school-age sample, its contribution was very small compared to that of the phoneme measures. In German, Wimmer, Landerl, and Schneider (1994) reported no predictive relations between early reading skills and rhyme awareness. In Spanish, Gonzalez, Gonzalez, Monzo, and Hernandez-Valle (2000) found that onset-rime units were largely irrelevant to reading.

In contrast, phoneme awareness has been found to be a robust and stable predictor of individual variation in reading both concurrently and longitudinally in most languages in which it has been investigated (Chiappe & Siegel, 1999; Comeau et al., 1999; Hoien et al., 1995; Mann & Foy, 2003; Marshall et al., 2001; Nation & Hulme, 1997; Oktay & Aktan, 2002; Porpodas, 1999). The predictive relations between phoneme awareness and reading development have been demonstrated using a large variety of tasks including phoneme counting, phoneme identification, phoneme segmentation, phoneme blending, phoneme deletion, phoneme substitution, and phoneme transposition. Overall, these results suggest that tasks measuring phoneme awareness may be more important in predicting reading development than tasks measuring larger unit processing, particularly in consistent orthographies.

Construct validity. A related question is whether phonological awareness represents a single ability, perhaps developing along a continuum of phonological skills, or whether it is composed of distinct phonological abilities. To put this question in operational terms, do tasks commonly used to measure phonological

awareness tap the same latent ability or do they measure independent cognitive abilities? A review of the developmental literature indicates that the main approach to this question has been based on the use of factor analysis.

Stahl and Murray (1994) used factor analysis to examine the structure of four phonological tasks varying in the type of operation required and in linguistic complexity. They identified a single factor, irrespective of whether they analyzed the data by the type of required operation or by the level of linguistic complexity. When the type of operation was used in the analysis, a single factor accounted for 73% of the variance; when the level of linguistic complexity was used, one factor accounted for 82% of the variance. Stahl and Murray argued that both dimensions can be used to define phonological awareness, with linguistic complexity accounting for somewhat more variance.

Stanovich et al. (1984) conducted an exploratory factor analysis on seven phoneme and three rhyming tasks with 49 preschool children and identified a single phonological factor, which received high loadings from all the phoneme—but not the rhyming—tasks. The one-factor model explained 48% of the variance in the test scores. The authors concluded that the different non-rhyming tasks were all measuring the same latent ability. Bowey (1996) also found a single factor underlying the performance of kindergarten children on two phonological tasks, a rhyme oddity task and a phoneme identity task; however, in contrast to the Stanovich et al. study, in her study, both the rhyme and the phoneme task loaded on the same factor. A possible criticism of Bowey's study is that the phonological

battery included only two tasks and that the phoneme task had low reliability (Cronbach alpha = .47), which casts some doubt on the appropriateness of this task for her sample.

Anthony and his colleagues (Anthony, Lonigan, Burgess, Driscoll, Philips, & Cantor, 2002) examined the relations among phonological awareness of syllables, that of rhymes, and that of phonemes in two preschool samples: two- and three-year-olds and four- and five-year-olds. Advanced confirmatory factor analysis techniques revealed that a one-factor model best explained the data from both groups of children. In a later study, Anthony and Lonigan (2004) re-examined data from four previous studies with a combined sample of more than 1,000 children ranging in age from two to seven years and found that the phonological awareness of the younger children was best explained with a one-factor solution, whereas the phonological awareness of the older children was best explained with a two-factor model that specified a separate rhyming factor. However, because the rhyming skills of these children were highly correlated with, and predictive of, their other phonological skills, the researchers concluded that phonological awareness is a single ability that manifests itself in a variety of skills during the preschool and early school years.

Schatschneider et al. (1999) examined the dimensionality of the phonological awareness construct using IRT. The researchers administered a battery of tests varying in task demands, the nature of target words (i.e., real words or nonwords), and the size of the unit of sound to be analyzed to 945 kindergarten,

first-, and second-grade children in order to assess the children's blending, segmentation, and sound categorization skills. They found evidence for a unitary construct underlying all the tasks they used.

Other researchers have proposed models with more than one factor. For example, Yopp (1988) conducted an exploratory factor analysis of ten phonological tasks with 96 preschool children and found two distinct factors underlying their performance—a large first factor with high loadings from the tests of phoneme segmentation, phoneme isolation, and phoneme counting as well as a smaller second factor with high loadings from phoneme deletion tasks. The larger factor accounted for 59% of the variance, and the smaller one for 9% of the variance in the test scores. Yopp called the first factor "simple phonemic awareness" and the second one "compound phonemic awareness" (p. 174) and argued that the deletion tasks tapped a different ability because they required more mental operations than did the other tasks and hence were more cognitively demanding. Hoien et al. (1995) found three separate factors in their study of Norwegian children: a syllable factor, a rhyming factor, and a phoneme factor. The scores on all three factors independently predicted reading ability.

Previous studies vary not only in the number of factors extracted but also in the very structure of these factors. For example, in Yopp's (1988) study, both extracted factors were phoneme-based, and rhyme tasks did not load on either of the extracted factors. In Stanovich et al.'s (1984) study, rhyme tasks were also irrelevant, although the researchers extracted only one factor. But in Muter, Hulme,

Snowling, and Taylor's (1997) study with four- to six-year-old English-speaking children, the two-factor solution included, in addition to a phoneme segmentation factor, a rhyme factor. However, in contrast to the results obtained by Yopp, in that study, the test of phoneme isolation and the test of phoneme deletion loaded on the same factor.

Why have results differed so widely? One possibility is that age and general developmental maturity may have affected performance. Many of the above studies included children of both preschool and school ages and it is possible that these groups of children were at different developmental stages and may not have been comparable. Indirect support for this possibility comes from observations that large-unit (e.g., rhyme) awareness and small-unit (e.g., phoneme) awareness do not develop simultaneously: Large-unit awareness appears to develop prior to reading acquisition, whereas small-unit awareness develops after reading exposure and is spurred by it (Ziegler & Goswami, 2005). For example, it has been shown that in order for phoneme awareness to develop, letter learning is required (see Share, 1995, for a review). There is also strong evidence for the reciprocal influence of phoneme awareness and reading on each other: Phoneme awareness influences reading development and reading influences the development of phoneme awareness (Perfetti, Beck, Bell, & Hughes, 1987; Share, 1995). In this sense, as Share has argued, phoneme awareness can be conceptualized as a reading skill rather than as a cognitive ability, much like letter-sound knowledge.

Many factor analytic studies have also been criticized for methodological as well as statistical shortcomings (e.g., see Anthony et al., 2002, for a review) including small sample sizes, exploratory procedures, low test reliability, floor and ceiling effects, and the use of single measures of phonological skills. It is also possible that some of the phonological measures, especially those requiring phoneme manipulation may not have been appropriate for some children as a certain degree of developmental maturity is required for a child to be able to perform the more cognitively demanding phonological tasks (Adams, 1990; Anthony & Lonigan, 2004). Moreover, factor analysis may not be an ideal tool for analyzing the structure of phonological awareness because, as has recently been shown in a series of simulation studies (e.g., Smith & Miao, 1994), it may not perform well in situations when the majority of test items are expected to load on just one factor.

Rapid Automatized Naming

Another correlate of reading ability is *phonological recoding in lexical access*, which refers to the automaticity with which a child can access a lexical referent for a written word by recoding the written symbols into a sound-based representation (Wagner & Torgesen, 1987). This skill is assessed using rapid automatized naming (RAN) tasks (Denckla & Rudel, 1976), in which a child is asked to name, under speeded conditions, a set of digits, colors, or pictures of highly familiar objects arranged randomly in a 10 x 5 array. This automaticity is

thought to be a function of both phonological processing and memory (Singleton, 2002).

RAN tasks have been used frequently in studies investigating the determinants of reading ability, and many researchers have reported a statistically significant relationship between RAN measures and early reading skills measured by both word reading accuracy and word reading speed (Gholamain & Geva, 1999; Holopainen et al., 2001; Kobayashi, Haynes, Macaruso, Hook, & Kato, 2005; Manis, Seidenberg, & Doi, 1999; Swanson et al., 2003; Wagner et al., 1997; Wimmer, 1993).

This relationship appears to be particularly strong in languages with transparent orthographies, where RAN measures have often been found to be a more robust predictor of reading performance than even phonological awareness measures. For example, Holopainen et al. (2001) found that in Finnish, naming speed in kindergarten was a significant predictor of reading speed in second grade. This was in contrast to phonological awareness measures (i.e., phoneme deletion and phoneme blending), which failed to differentiate early from late decoders. De Jong and van der Leij (2002) examined the effect of phonological abilities on reading development in a longitudinal study of 140 Dutch children and found that the relation of word-reading speed with rapid naming was consistently stronger than the relation of word-reading speed with phonological awareness, although the effects of both RAN and phonological awareness were limited to the first grade.

Szczerbinski (2003) reported that RAN was a strong and independent predictor of reading ability in an unselected group of Polish elementary school children.

The strong relationship observed between RAN measures and reading in transparent orthographies may reflect the special importance of rapid naming to the development of word-reading speed as opposed to word-reading accuracy (Manis et al., 1999; Torgesen et al., 1997; Wimmer, 1993; Wolf & Bowers, 1999). It has long been observed that in languages with transparent orthographies, where the relationship between phonemes and graphemes is straightforward, even readers with phonological deficits have little trouble acquiring decoding accuracy (Wimmer, 1996; Wimmer et al., 2000), and it is the speed of decoding, rather than the accuracy that appears to be the distinguishing feature between good and poor readers (de Jong & van der Leij, 2003; Landerl, 2003; Lyytinen et al., 2004). Because rapid naming is more closely related to reading speed than to reading accuracy, individual differences in rapid naming should have a greater influence on reading development in languages where impaired reading speed, rather than accuracy, is the hallmark of reading disability.

In fact, rapid naming problems seem to be one of the main characteristics of dyslexic children learning to read in a transparent orthography. In Dutch, deficits in RAN in dyslexic readers were found to persist largely unchanged from kindergarten into adolescence, whereas deficits in phonological awareness were developmentally limited and disappeared by sixth grade (de Jong & van der Leij,

2003). In German, rapid naming tasks have been found to be the best differentiators between dyslexic and normal readers (Landerl, 2003).

Recent research has also suggested that rapid naming may have different relationships with reading in average versus poor readers. Several researchers have found that rapid naming is more predictive of reading ability in readers with dyslexia than in average readers (Cardoso-Martins & Pennington, 2004; McBride-Chang & Manis, 1996; Meyer, Wood, Hart, & Felton, 1998). For example, Cardoso-Martins and Pennington examined the contribution of rapid naming as well as phonemic awareness to early literacy acquisition in children at high and at low familial risk of dyslexia. They found that rapid naming contributed unique variance to reading and spelling ability only in the high-risk group, whereas phonemic awareness was a strong predictor of reading in both groups of children. Further, the predictive power of rapid naming in average readers appears to be limited to an early period of development and is often not found with mature readers (Torgesen et al., 1997; Wagner et al., 1997), whereas in dyslexic populations, rapid naming remains a reliable predictor of reading ability throughout elementary and junior high school years (Meyer et al., 1998; Wolf, Bowers, & Biddle, 2000).

Similarly, the patterns of association between rapid naming and other cognitive abilities on the one hand and reading on the other appear to differ depending on the skill of the reader. For example, McBride-Chang and Manis (1996) used structural equation modeling to investigate patterns of association

between reading, phonological awareness, verbal intelligence, and rapid naming in average and poor readers. They found that for poor readers, both rapid naming and phonological awareness were associated with word reading, whereas for average readers, phonological awareness and verbal intelligence were significantly associated with reading, but rapid naming was not.

How should the relationship between rapid automatized naming and reading be interpreted? There are two hypotheses. One (e.g., Share, 1995; Wagner and Torgesen, 1987) views rapid naming as a phonological ability because it involves accessing and retrieving phonological information from long-term memory. For example, Torgesen et al. (1997) have argued that RAN tasks "have sufficient face and predictive validity as phonological measures to include them in the 'phonological family'" (p. 162). In this view, the relationship between reading and rapid naming is similar to that between reading and other phonological abilities, and deficits in rapid naming are interpreted as part of a more global phonological deficit that impedes efficient access to, and retrieval of, phonological representations (Landerl, 2003; Snowling, 2000, p. 44).

An alternative interpretation (Wolf & Bowers, 1999; Wolf et al., 2000) emphasizes both linguistic and visual processes involved in RAN and explains RAN deficits as an impaired timing mechanism. For example, Wolf and her colleagues have argued that although rapid naming, like any linguistic task, involves accessing phonological codes, it consists of *multiple* processes, many of which are nonphonological (i.e., attentional, perceptual, semantic, and articulatory).

These processes, together with the precise timing requirements and serial processing involved in RAN, make RAN tasks different from purely phonological tasks. In this view, children with rapid naming deficits cannot quickly process visual and phonological codes for printed letters and as a result, cannot form orthographic representations of words as efficiently as normal readers.

Whatever the nature of RAN, its contribution to reading ability appears to be independent from the effects of phonological awareness and other cognitive abilities, such as verbal short-term memory. Support for this conclusion comes from studies demonstrating that RAN adds independent variance to reading ability over and above that added by phonological awareness (Manis et al., 1999; McBride-Chang & Manis, 1996); from studies showing that RAN and phonological awareness measures correlate only weakly (Manis et al., 1999; Wimmer et al., 2000); and from studies showing that RAN and phonological awareness tasks load onto separate factors (Gottardo, 2002; Van den Bos, 1998). Additional evidence comes from studies that have identified three groups of poor readers: those with a single deficit in rapid naming, those with a single deficit in phonological awareness, and those with a double deficit in both naming speed and phonological awareness (e.g., Wimmer et al., 2000; Wolf & Bowers, 1999).

In summary, the following conclusions can be drawn regarding the relationship between RAN and reading. First, RAN appears to contribute unique variance to reading acquisition and is most probably distinct from other phonological measures. Second, the contribution of RAN to reading ability seems

larger in younger, less mature, and more disabled readers. Third, the relationship between RAN and reading may differ depending on the orthography, with more transparent orthographies showing a stronger relationship. Finally, while RAN measures account for variance in both word-reading accuracy and word-reading speed, they are more strongly related to speed measures than to accuracy measures.

Verbal Short-Term Memory

Another correlate of reading ability is verbal short-term memory (STM), which refers to the temporary storage of verbal information (Wagner & Torgesen, 1987). When verbal STM is discussed from the point of view of its function rather than structure, it is referred to as *verbal working memory*, a term that emphasizes its performance, or the ability to store and process verbal input. The capacity of verbal STM is limited to about seven to nine items (Aaron & Joshi, 1992), while its duration is only about 2 seconds (Baddeley, 1992). It is further hypothesized (e. g., Baddeley) that verbal information is retained in working memory by continuous repetition inside a phonological loop to prevent it from fading. In reading, this sub-vocal repetition allows verbal information to be retained for a period of time long enough to synthesize individual speech sounds into syllables and words.

It is generally accepted in the reading literature that verbal information is held in short-term memory in phonological form (e.g., Aaron & Joshi, 1992; Snowling, 2000). Therefore, tasks measuring verbal STM performance also reflect the ability to convert verbal input into its phonological form, or to represent

mentally the phonological features of language (Torgesen, Wagner, & Rashotte, 1994). For this reason, verbal STM has often been included in phonological skills, along with phonological awareness and rapid naming (e.g., Wagner & Torgesen, 1987). Furthermore, many researchers (e.g., Snowling, 2000; Turner, 1997) treat poor performance on verbal STM tasks as an indication of phonological deficits. Accordingly, children showing problems with converting verbal material into its phonological form would be expected to encounter difficulty in storing and retaining information in verbal STM.

Common tasks tapping verbal STM include word repetition, nonword repetition, and digit span. Performance on these tasks depends on the child's ability to store information in phonological form (Share, 1995). It has been argued that digits are a purer form of verbal STM capacity than words because the influence of prior vocabulary in digit repetition is lower, and the memory component is higher (Turner, 1997). Nonwords are even better, especially at younger ages, because, unlike words, nonwords are not prone to lexicality effects: They do not have lexical entries in long-term memory and their repetition cannot be facilitated by long-term memory.

Performance on verbal STM tasks is highly associated with both current and later reading development (e.g., Turner, 1997), although the nature of this relationship appears to be correlational rather than causal. For example, in Nation and Hulme's (1997) study of precursors of reading in English-speaking elementary school children, verbal STM performance (measured by digit span) predicted,

together with age, 60% of the variance associated with reading scores. Verbal STM has also been shown to be related to performance on phonological awareness tasks (Hansen & Bowey, 1994). Some researchers (Singleton, 2002) have even argued that of all psychological measures, verbal STM may be the most robust source of difference between good and poor readers.

Furthermore, verbal STM appears to contribute variance to reading achievement over and above that contributed by other phonological skills. For example, Passenger, Stuart, and Terrell (2000) followed a group of 80 preliterate children during their first year of formal schooling in order to investigate the relationship between phonological awareness and verbal short-term memory (measured by digit span and nonword repetition) in reading development. They found that both phonological awareness and verbal STM made significant and distinct contributions to early reading skills. Similar results were reported by Hansen and Bowey (1994), who found that verbal STM accounted for unique variation in reading development of second-grade children, over and above that accounted for by phonological awareness.

Deficits in verbal STM have been shown to be a strong indicator of difficulties with reading acquisition (Gang & Siegel, 2002; Nelson & Warrington, 1980). In a longitudinal study of reading precursors, Mann and Liberman (1984) found that inferior performance on verbal STM tasks in kindergarten was predictive of later reading problems. Bowers, Steffy, and Tate (1988) found that in dyslexic children, verbal STM was a good predictor of reading ability even when nonverbal

intelligence was statistically controlled. In fact, deficits in verbal STM in dyslexia are so pervasive that Turner (1997) has called them "the hallmark of dyslexia" (p. 86), while Snowling (2000) has referred to them as "the most consistently reported area of difficulty" for dyslexics (p. 35). These deficits can be detected even in adulthood and even in fully compensated dyslexics (e.g., Paulesu et al., 1996), and they appear to be specific to tasks requiring phonological coding (Share, 1995).

Cross-language research has largely supported the above findings regarding the role of verbal STM in reading development and reading disability. STM deficits in dyslexia have been observed in languages as diverse orthographically as Arabic, Finnish, Greek, and Hebrew (see Smythe et al., 2004, for a review). For example, in Greek, Porpodas (1999) compared good and poor readers and found that in addition to differences on phonological tasks, the groups differed significantly on most short-term memory tasks. Gholamain and Geva (1999) found evidence of verbal memory involvement in the development of reading skills in the two languages of their bilingual English-Persian subjects despite considerable orthographic differences between English and Persian. In Dutch, verbal STM measured by a digit span test contributed significantly to the prediction of dyslexia (Elbro et al., 1998). In Hebrew (Share, 2003), verbal STM assessed in kindergarten accounted for 16% of the variance in word decoding ability a year later, making a significant and unique contribution to the prediction of first-grade decoding.

But how, theoretically, could deficits in verbal STM cause problems in learning to read? There are two alternative interpretations. One centers on the

independent contribution of verbal STM to individual differences in reading and ascribes dyslexic children's difficulties to their general lack of capacity for the simultaneous processing and storage of information (de Jong, 1998). In this interpretation, memory processes are accorded a direct role in learning to read and problems in this area are thought to be memory-specific. For example, Tractenberg (2002) suggested a conceptualization of reading achievement as based on both memory and phonological awareness to account for the range in severity of reading disability. In a study comparing the reading performance of adults with reading disability (RD) and those with hearing impairment, she found that while both groups showed deficits in phonological awareness, only the RD group demonstrated deficits in verbal STM. Moreover, adults with hearing problems achieved significantly higher reading levels than did adults with RD in spite of their similar phonological segmentation skills. Based on these results, Tractenberg argued for a separable influence of verbal STM on reading. She further argued that average reading achievement was possible in the absence of phonological awareness but was unlikely without intact memory.

In an alternative interpretation (e.g., Hatcher & Snowling, 2002; Snowling, 2000), limitations in verbal STM displayed by dyslexics are taken as a sign of their difficulties at the level of phonological representations, or representations of the phonological forms of words. These representations are thought to be poorly specified and/or incomplete in dyslexic children, placing a heavy burden on the children's cognitive resources and restricting the amount of information dyslexic

children can hold in short-term memory. This impairment affects learning to read by reducing the child's ability to blend the sounds of visually presented letters to form words during decoding.

Whatever the nature of the memory deficit, limitations in verbal STM have far-reaching consequences for the child's linguistic development. Gathercole and her colleagues (Gathercole & Baddeley, 1989; Gathercole, Hitch, Service, & Martin, 1997) have shown, for example, that in young children, short-term memory is critical to vocabulary acquisition. In a longitudinal study of children at familial risk for dyslexia, Scarborough (1990) found that at three years of age, children who later developed dyslexia—which in English is often characterized by STM deficits (Snowling, 2000; Turner, 1997)—showed deficits in receptive vocabulary and object-naming ability. Children with memory deficits would also be expected to be considerably slower than normally developing children in learning foreign languages (Johansson, 2006) because it takes them much longer to store new phoneme sequences in memory. This claim has recently found some empirical support (Crombie, 1997; see also Ganschow, Sparks, & Javorsky, 1998; Ganschow, Sparks, & Schneider, 1995; Robertson, 2000, for reviews).

Reading Comprehension

The Nature of Reading Comprehension

In the previous sections, I have reviewed research on determinants of early word recognition and the role of orthography in this process. However, there is

much more to reading than the ability to recognize words—the reader must also understand what he or she is reading. Reading comprehension is a complex activity that involves, besides graphophonemic decoding, lexical access and retrieval, syntactic parsing, and higher-order subprocesses that ensure the building of a coherent mental representation of the text (Megherbi et al., 2006, p. 136). It is often described as a hierarchically layered process (e.g., Adams, 1990): At the bottom level, the reader must decode each individual word and retrieve each word's meaning; at the next level, the reader must parse the sentence into syntactically acceptable units and assemble the meanings of the words in each unit to form a "composite interpretation" (p. 141); finally, the reader must interpret the just-read phrase taking into account the context and the overall meaning of the text.

Thus, at lower levels, successful reading performance will depend on the accuracy and speed of word recognition as well as the ease with which the reader can retrieve lexical meanings of words (Adams, 1990; Vellutino et al., 2004). Inadequate facility in word recognition—either in terms of fluency or accuracy—will impede reading comprehension even in readers with adequate oral language comprehension. At higher levels, successful reading performance will depend on the reader's ability to interpret facts, understand relationships, and use contextual information and his or her own background knowledge. Comprehension at this level is not automatic and may require "retrieval of particular facts or events presented many pages earlier in the text[,]...consideration of knowledge and construction of argument that are entirely extraneous to the text...[and] the critical

and inferential activities necessary for putting... information together...[as well as] active attention and thought" (Adams, p. 142). Recent research has shown that the two components of reading, decoding and comprehension, can be clearly differentiated and that they have at least partly different determinants (de Jong & van der Leij, 2002; Shankweiler et al., 1999).

Reading comprehension may require a number of different skills, depending on the reader's experience, the purpose for reading, and other factors. In his review of several taxonomies of reading comprehension skills, Van Etten (1978) described two types of skill that are particularly important to early reading development—literal comprehension, which refers to the ability to comprehend facts and literal meanings and repeat them from a text, and interpretation, which involves understanding unstated meaning and relies on the reader's ability to make generalizations and discover relationships. Following Van Etten, I use the word "comprehension" to refer to both the ability to comprehend facts and the ability to interpret them. In the next sections, I will review previous research on the determinants and correlates of reading comprehension.

Components of Reading Comprehension: The Simple View of Reading

Factors related to reading comprehension can be roughly divided into two categories—those related to skills required for word recognition and those related to skills required for extracting word meanings and integrating them into a holistic interpretation of a larger text. Thus, reading comprehension can be conceptualized

as a function of decoding and linguistic comprehension. This is the Simple View of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), according to which R = D x C, where R is reading comprehension, D is decoding defined as "efficient word recognition" (Hoover & Gough, p. 130), and C is linguistic comprehension defined as "the ability to take lexical information…and derive sentence and discourse interpretations" (Hoover & Gough, p. 131).[5] Recently, it has been shown that an additive formula, where R = D + C, is just as effective in explaining variance in reading comprehension (Joshi & Aaron, 2000).

The Simple View formula postulates an independent effect on reading comprehension for both word decoding ability and linguistic comprehension ability, and it predicts impaired reading ability in a child with deficient decoding skills and/or deficient linguistic comprehension ability. Several lines of evidence support the separability of decoding and linguistic comprehension as well as the predictive validity of this formula.

The first one comes from correlational and predictive studies. Hoover and Gough (1990) confirmed the validity of the Simple View formula in a sample of more than 500 bilingual English-Spanish children in the first through fourth grades. In their study, the largest correlations with reading comprehension were of the product of decoding and listening comprehension, which ranged from .84 in first grade to .91 in third and fourth grades; the next largest values were for decoding

---

[5] Hoover and Gough (1990) operationalized linguistic comprehension as listening comprehension. However, tests of vocabulary and sentence processing have also been used in previous research to measure linguistic comprehension.

ability in the first two grades (r = .80-.84) and for listening comprehension ability in the last two grades (r = .80-.87). Joshi and Aaron (2000) found that the product of decoding and listening comprehension explained about 48% of the variance in reading comprehension in third-grade children. Cutting and Scarborough (2006) reported that both decoding and oral language skills (measured by tests of vocabulary and sentence processing) made unique contributions to the prediction of reading comprehension, accounting for 49-72% of the total variance, depending on the reading comprehension measure used.

The second line of evidence comes from longitudinal studies. Nation and Snowling (2004) examined the relationship between oral language skills and both reading comprehension and word recognition in a sample of 72 average readers who were followed from the age of 8.5 for more than five years. Oral language measures included a test of expressive vocabulary, a test of listening comprehension, and a test of semantic fluency, in which a child is given a semantic category and asked to produce, within a time limit, as many items for that category as possible. The three measures of oral language accounted for a significant portion of unique variance in reading comprehension (between 4% and 14%) both concurrently and longitudinally, even when the effects of age, IQ, phonological skills, and prior reading comprehension were taken into account. These oral language measures were also significant predictors of word recognition, both concurrently and longitudinally, although the portion of unique variance that they explained was small (i.e., 2-4%). In another longitudinal study, de Jong and van der

Leij (2002) examined the effects of decoding ability and linguistic comprehension (assessed with tests of receptive and productive vocabulary and listening comprehension) on the development of reading comprehension in a sample of 140 Dutch children. Both word decoding and linguistic comprehension measured in the first grade predicted third-grade reading comprehension even when autoregressive effects of reading comprehension were taken into account.

Factor-analytic evidence further supports the separability of decoding and linguistic comprehension constructs, at least in some readers. In a study of poor teenage readers with low to average verbal and listening ability, Savage (2006) found two distinct factors with eigenvalues exceeding 1 underlying reading ability: a decoding factor, with strong loadings from word and nonword reading and spelling tests, and a comprehension factor, with loadings from tests requiring higher order verbal processing including verbal ability, listening comprehension, and writing. The decoding factor explained about 48% of the variance and the comprehension factor about 18%. As predicted by the Simple View, the only variable that loaded strongly on both decoding and comprehension factors was reading comprehension.

Furthermore, models based on the Simple View have been shown to allow a fairly accurate and relatively stable classification of readers into subgroups depending on their profile of strengths and weaknesses. For example, using *z* scores of less than -1 as cut-off values for poor performance in word recognition or listening comprehension, Catts, Hogan, and Fey (2003) subgrouped their sample of

poor readers into four groups—those with adequate listening comprehension but poor word recognition, those with poor performance in both word recognition and listening comprehension, those with relative strengths in word recognition but poor listening comprehension, and those with adequate performance in both word recognition and listening comprehension. The subgroups differed significantly from one another in reading comprehension as well as in nonverbal IQ, although they did not represent homogeneous clusters with clear boundaries. This classification remained relatively stable from second to fourth grade.

In another study, Shankweiler et al. (1999) used a cut-off score procedure with a buffer zone to classify discrepant readers into categories. They identified two groups of readers—those with poor reading comprehension but normal decoding and those with poor decoding but normal reading comprehension, although both skills in both groups fell short of the same skills observed in unimpaired children. In other words, the discrepant readers were impaired in both decoding and reading comprehension relative to normal children, although they had relative strengths that distinguished them from those readers who were uniformly low on both decoding and reading comprehension. Overall, this and similar studies show that it is possible to differentiate between readers with specific weaknesses in reading comprehension and those with specific weaknesses in decoding.

The Role of Decoding

There is a strong and positive connection between fluent word recognition and reading comprehension (Berninger, Abbott, Vermeulen, & Fulton, 2006; Rupley, Willson, & Nichols, 1998). Fluent word recognition manifested in accurate and rapid decoding of print permits the reader to construct the meaning of text (Pikulski & Chard, 2005, p. 510), whereas slow word recognition directly affects comprehension "by inhibiting [the] chunking of information into meaningful information units" (Rupley et al., 1998, p. 155). In English, correlations between reading comprehension and decoding (measured by real word reading) are often in the .60-.70 range (Swanson et al., 2003) and exceed those between reading comprehension and verbal ability, IQ, or memory. Slow and inaccurate decoding has been shown to be the best predictor of deficits in reading comprehension (e.g., see Lyon, 1996, for a summary of research in the U.S.).

How does decoding affect reading comprehension? It is generally assumed (Adams, 1990; Oney & Goldman, 1984; Stanovich, 1981; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Korne, 2003) that both lower order processing required in identifying letters and assembling word pronunciations and higher order processing required in reading comprehension share a limited-capacity cognitive system. The automaticity of lower order processing (e.g., decoding) frees up memory and other cognitive resources so that they can be diverted to higher order processing required in text comprehension. In contrast, slow letter identification and serial, letter-by-letter assembly of word pronunciations places huge demands

on memory and other cognitive resources and prevents the system from making more capacity available for the higher order processes needed in text comprehension. Thus, it is not so much the accuracy as the speed of word decoding that is expected to affect reading comprehension. As Adams has put it, "[t]he greater the time and effort a reader must invest in each individual word, the slimmer the likelihood that preceding words of the phrase will be remembered when it is time to put them all together" (p. 141).

Many studies have demonstrated that reading speed directly affects comprehension, accounting for additional variance in reading comprehension ability (Berninger et al., 2006; Cutting & Scarborough, 2006; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Rupley et al., 1998). For example, in a study of 97 English-speaking first through tenth graders, Cutting and Scarborough (2006) found that reading speed explained an additional 1-6% (depending on the test used to operationalize reading comprehension) of the variance over and above that accounted for by reading accuracy and oral language skills (assessed with tests of lexical knowledge and sentence processing). In another study, Rupley et al. (1998) showed through structural equation modeling that young children's reading comprehension was directly influenced by their rate of processing text: For children in the first and second grades, there was a significant path from reading rate to word recognition accuracy and from word recognition accuracy to reading comprehension.

Reading speed may be particularly relevant to reading comprehension in languages with more transparent orthographies than English. For example, de Jong and van der Leij (2002) found that in Dutch, word decoding speed influenced the development of reading comprehension from first to third grade. Correlations between word decoding speed and reading comprehension in the first and third grades were .70 and .61, respectively, indicating that this relationship remained stable over the observed period. Oney and Goldman (1984) compared Turkish and American first and third graders on nonword reading (scored for both accuracy and speed) and passage comprehension. Correlations between the children's mean decoding speed for correctly pronounced nonwords and the number of correct responses on the reading comprehension test were moderate and significant for both groups of Turkish children. Correlations for the American first graders were marginally significant, and those for the American third graders were not significant. These results indicate that for languages with a higher degree of letter-sound correspondence, greater automaticity with decoding may be associated with better comprehension.

Finally, the relationship between decoding and reading comprehension may depend on how reading comprehension is operationalized, as has been shown by Berninger and her colleagues (2006). The researchers administered several tests of reading comprehension including a cloze, a meaning judgment test, an open-ended question test, and a multiple-choice test to a group of poor readers in the fall and in the spring of second grade. They found that the scores on each test were uniquely

explained by a different predictor, and that this also depended on the time of testing. For example, in the fall, meaning judgment was uniquely explained by word reading rate, whereas cloze comprehension was uniquely explained by word reading accuracy; multiple choice and open-ended comprehension were uniquely explained by both accuracy and rate of word reading. In the spring, multiple-choice comprehension was uniquely explained by word reading accuracy, whereas cloze and open-ended comprehension were uniquely explained by both accuracy and rate of word reading. However, in a confirmatory factor analysis, all reading comprehension measures loaded on a single factor, indicating that they all measured essentially the same thing. This factor did not change throughout the second grade.

Differences in the operationalization of key concepts (e.g., reading comprehension and decoding) may also explain why some researchers have failed to find any relationship between decoding and reading comprehension. For example, Goff, Pratt, and Ong (2005) found that when measured by passage reading (as opposed to the more conventional word list reading), reading speed did not make an independent contribution to reading comprehension.

The Relationship Between Decoding and Listening Comprehension

There is evidence to suggest that the relative importance of decoding versus listening comprehension in predicting reading comprehension may depend on the developmental stage of the child. Facility in decoding appears to carry much

greater weight as a determinant of reading comprehension in children at the early stages of reading development than in children at later stages, whereas listening comprehension becomes more important at later stages of reading development (Vellutino et al., 2004). This "developmental asymmetry" (p. 6) is evident in the strength of correlations between decoding and reading comprehension and between listening comprehension and reading comprehension. For example, in Hoover and Gough's (1990) summary of correlational research, decoding correlated with reading comprehension to a greater degree than did listening comprehension in lower grades (r = .55 for decoding and .35 for listening comprehension), whereas listening comprehension correlated with reading comprehension to a greater degree than did decoding in higher grades, when readers became more experienced. Hoover and Gough's study found further support for this conclusion.

Studies investigating reading comprehension in less skilled beginning readers have also demonstrated a greater contribution to reading comprehension from decoding than from listening comprehension. For example, in a study of poor readers, Shankweiler et al. (1999) reported very high correlations between reading comprehension and word and nonword reading (r = .89 and .79, respectively) but only moderate correlations between listening comprehension and reading comprehension (r = . 58). To assess more accurately the contributions of decoding and listening comprehension to reading comprehension, the researchers partialed out of the correlation the contribution of the overlapping measure. The correlation of decoding with reading comprehension partialed on listening comprehension was

.65, whereas the correlation of listening comprehension with reading comprehension partialed on decoding was substantially weaker, .37. These results indicate that in less skilled readers, tests evaluating decoding may be better predictors of reading comprehension performance than tests evaluating listening comprehension.

The relative weight of decoding versus listening comprehension in determining reading comprehension may also depend on the linguistic features of the language, including orthography. Megherbi et al. (2006) assessed the contribution of decoding skill and listening comprehension to reading comprehension in French first and second graders. They found that reading comprehension correlated with listening comprehension much more strongly than it did with decoding in both grades. Hierarchical regression analyses showed that in first grade, listening comprehension added 29% to the proportion of variance explained by decoding, while decoding added 17% to the proportion of variance explained by listening comprehension; in second grade, listening comprehension added 36% of the variance over and above that contributed by decoding, while decoding added 8% to the proportion of variance explained by listening comprehension. Thus, although both listening comprehension and decoding were significant predictors of reading comprehension, the proportion of variance uniquely explained by listening comprehension was substantially greater in both grades. This result differs from those typically reported for English, where in early grades, decoding has been shown to carry more weight than listening

comprehension in determining reading comprehension (e.g., Hoover & Gough, 1990). Megherbi and her colleagues argued that the transparency of the language may affect the weight of abilities that contribute to reading acquisition.

Finally, the relationship between reading comprehension and listening comprehension appears to change with development. For example, Diakidoy, Stylianou, Karefillidou, and Papageorgiou (2005) examined the relationship between listening and reading comprehension in a sample of 612 Greek-speaking children in four different grades. Both reading and listening comprehension tests required the verification of literal and inferential statements, and all required a yes/no answer. Listening and reading comprehension scores were significantly correlated with each other at all grade levels and their relationship became progressively stronger in higher grades (e.g., $r = .44$ in Grade 2 vs. $r = .63$ in all the other grades). These results are consistent with the correlational evidence obtained in English (see above) demonstrating a progressively stronger relationship between reading comprehension and listening comprehension as children become increasingly more skilled at word decoding. Diakidoy et al. also found significant differences in the reading versus listening comprehension performance between the grades: In the second grade, the children's listening comprehension performance exceeded their reading comprehension performance; in the fourth and sixth grades, the children's listening comprehension performance was comparable to their reading comprehension performance, and in the eighth grade, their reading

comprehension performance became better than their listening comprehension performance.

Verbal Skills

Many researchers have argued that progress in reading beyond the basic stages depends to a large degree on verbal skills, primarily on vocabulary (Nation & Snowling, 2004; Pikulski & Chad, 2005). Torgesen et al. (1997) found substantial effects of vocabulary knowledge on reading comprehension in second- to fifth-grade children. The influence of vocabulary on the development of reading comprehension appeared to increase with the grade level: Second-grade vocabulary explained 24% of the variance in fourth-grade reading comprehension, whereas third-grade vocabulary explained 43% of the variance in fifth-grade reading comprehension. However, when the autoregressive effects of prior reading were taken into account, these percentages dropped to 2% and 10%, respectively. In a study of independent predictors of reading comprehension, Goff et al. (2005) found that receptive vocabulary explained about 4% of the variance in reading comprehension after irregular word reading was statistically controlled. There is also strong correlational evidence for the connection between vocabulary and reading comprehension. For example, in a summary of research on reading comprehension conducted in 15 countries, Aaron and Joshi (1992) cited correlations between vocabulary and reading comprehension ranging from .66 to .75.

One explanation for the positive link between vocabulary and reading comprehension was proposed by Ehri (1995, 2002), who argued that in order to be activated, the phonological and meaning aspects of the words one reads must be present in one's mental lexicon. In this sense, fluent word recognition depends on the reader's familiarity with the words' pronunciation and their meaning, or, in other words, on the reader's vocabulary knowledge.

Deficits in reading comprehension have been associated with weaknesses in vocabulary (Nation & Snowling, 1998, 1999) and related semantic skills. Nation and Snowling (1998) compared poor comprehenders to normal readers matched for decoding skills and IQ. Poor comprehenders had weak semantic processing abilities and experienced greater difficulty reading low-frequency and exception words. The researchers argued that individual differences in semantic processing may underlie individual differences in reading comprehension and speculated that these weaknesses may be causally related to poor readers' comprehension problems. There was also evidence that semantic difficulties constrained the development of effective word recognition skills in poor comprehenders, leading to problems with irregular and low-frequency word reading.

There is some evidence, however, that the role of vocabulary in reading comprehension is smaller than the role of listening comprehension, at least for some readers. For example, in a longitudinal study of the determinants of reading comprehension in Dutch, de Jong and van der Leij (2002) found that the effect of listening comprehension on reading comprehension was greater than that of

vocabulary. Savage (2001) assessed the role of listening comprehension and verbal cognitive skills in predicting reading comprehension in a group of English-speaking poor readers. The readers' decoding skills were kept constant. The best predictor of reading comprehension was listening comprehension. Although reading comprehension correlated significantly with verbal ability, this correlation was substantially weaker than the correlation with listening comprehension. A possible criticism of this study is that it involved only 14 mixed-ability teenagers referred for assessment due to a demonstrated discrepancy between cognitive and reading skills.

The Role of General Intelligence

There have been a number of empirical studies investigating the relationship between general intellectual ability (as measured by various IQ tests) and reading achievement defined both as single word reading and as reading comprehension in school children. Many researchers have reported what Elifson, Runyon, and Haber (1998) call *moderate* correlations. For example, Stanovich et al. (1984) summarized several decades of research on the IQ-reading relationship and showed correlations for second-grade children ranging from .28 to .76, with the mean correlation of .49 and the median of .48. They went on to investigate the relationship between IQ and reading comprehension in a sample of first, third, and fifth graders using two intelligence measures, *Raven's Progressive Matrices* (RPM), which was used to assess nonverbal ability, and the *Peabody Picture*

Vocabulary Test (PPVT), which was used to assess verbal performance. Both the RPM and the PPVT had moderate correlations with reading comprehension (r = .33 and .34, respectively, for first graders; r = .42 and .59 for third graders; and r = .56 and .58 for fifth graders). Stanovich et al.'s research, however, was based on a small sample–56 children in the first grade, 18 in the third grade, and 20 in the fifth grade.

Studies based on larger samples have consistently yielded more stable correlations. Particularly noteworthy in this respect are the studies conducted by Naglieri (1996) and Naglieri and Ronning (2000), who used large samples that were representative of the U.S. population on a number of key demographic variables. For second- through ninth-grade students, (N = 2,125), Naglieri reported correlations ranging from .43 to .58 between a test of nonverbal ability and a reading test that combined measures of decoding and reading comprehension; for a representative sample of more than 22,000 preschool through twelfth-grade students, Naglieri and Ronning reported a median correlation of .52 between nonverbal IQ and word reading and that of .56 between nonverbal IQ and reading comprehension. In both studies, the correlations were greater in higher grades. These studies have high ecological validity because they employed nonverbal measures of intelligence and excluded verbally loaded, reading-related subtests in order to avoid any contamination of the results due to content overlap. The use of nonverbal rather than verbal or composite measures of IQ may be particularly justified with younger children because reading skills in the first years of literacy

acquisition appear to relate more closely to perceptual and pattern-recognition skills, rather than to overtly linguistic skills such as vocabulary (Turner, 1997).

In a recent study, Tiu, Thompson, and Lewis (2003) reported a correlation of .60 between nonverbal IQ as measured by the *Wechsler Intelligence Scale for Children, Third Edition,* and reading comprehension. They conducted a series of hierarchical regression analyses and found that nonverbal ability was a significant predictor of reading comprehension and that it added more to the prediction of reading comprehension than either listening comprehension or processing speed. The authors also compared the relationship between IQ, listening comprehension, processing speed, decoding, and reading comprehension in a reading-disabled and a non-disabled group. They found that while IQ was a significant predictor of reading comprehension, processing speed, and listening comprehension in both groups, in the reading-disabled group, it also predicted decoding. This last result implies a greater importance of nonverbal intelligence in predicting reading ability in reading-disabled children.

Comparisons of correlations between reading and IQ and reading and other cognitive and processing abilities reveal striking similarities. Swanson et al. (2003) meta-analyzed 35 studies in order to examine correlations between various cognitive and processing abilities on the one hand and both word recognition and reading comprehension on the other. They found that the magnitude of the correlations between word reading and such abilities as phonological awareness, rapid automatized naming, vocabulary, and memory was in the same range as that

of the correlations between word reading and IQ (i.e., around .40). The correlation between IQ and reading comprehension was even higher, with $r = .65$. Similar findings were reported by Hammill (2004), who analyzed the combined results of three meta-analyses examining the validity of a variety of ability measures in predicting reading achievement both concurrently and longitudinally. In all, Hammill reviewed 452 studies and analyzed 10,754 coefficients to identify 10 ability clusters. His results showed that all non-print-related abilities (e.g., phonological awareness, RAN, and memory) correlated with reading to the same degree as did intelligence, a finding suggesting that all these abilities may be components of general developmental maturity and may simply represent different aspects of general intelligence. This conclusion would be consistent with Turner's (1997) observation that "most reading problems are caused by low intelligence" (p. 37).

In summary, it appears that in young children, IQ is at least moderately correlated with both decoding skills and reading comprehension and that the magnitude of these correlations is comparable to that of correlations between reading and other processing and cognitive abilities. It is important to acknowledge, however, the rather controversial nature of the concept of IQ as well as the widely-held belief in the field that no current IQ test can accurately measure intelligence as intellectual potential or predict future achievement in any valid sense (Fletcher, Coulter, Reschly, & Vaughn, 2004; Stanovich, 1991). Thus, the term IQ (or *nonverbal ability*) as discussed in this study does not refer to intelligence as

intellectual potential or aptitude for learning; rather, IQ test scores are used here as a "gross measure of current cognitive functioning" (Stanovich, p. 10), which to some degree reflects past learning.

Teaching Reading in Russian Schools

Education in Russia is compulsory and free of charge to all school-age children. Russian parents have the option of sending their children to school at age six or seven and many send their children to preschool until age seven.

According to the guidelines adopted by the Russian Ministry of Education for Russian language arts instruction in mainstream public schools (Alekseenko et al., 2003), the elementary school program in Russian language arts includes three elements: literacy instruction, literature reading, and instruction in the formal structure of the Russian language. The goal is to develop children's speech and to teach children to read fluently and meaningfully. Emphasis is placed on achieving high reading speed without enunciating the text. This is achieved by teaching children to appreciate letter-sound correspondences; to read smoothly, first syllable-by-syllable, then in whole words; and to understand rules of grammar. Literacy is taught for the first six to eight months of first grade and is then replaced by Russian language arts, which is taught through ninth grade. Reading is taught in a separate class, from first to fourth grade, and is then replaced by literature and composition.

The traditional course in literacy is phonics based (Grigorenko, 2003). Children receive instruction in the practical and theoretical aspects of phonetics, morphology, orthography, and grammar and acquire specific skills that enable them to decode words. Phonetics is taught through explanations of such concepts as sounds, speech organs, vowels, consonants, voiced versus voiceless sounds, hard versus soft sounds, stress position, and letter-sound correspondence. These concepts are taught explicitly and are covered thoroughly in textbooks. Exercises aimed at developing metalinguistic skills such as breaking up a word into component sounds or recognizing letters and the sound each letter represents are an integral part of literacy instruction. Morphology is taught by explaining parts of speech (e.g., nouns, verbs), word parts (e.g., prefixes, suffixes), and rules for word building. Orthography is taught through the introduction of the main principles of Russian orthography and rules of grammar.

Like all other classes, Russian language and reading classes last 45 minutes and usually begin with the teacher instructing children on what to do. According to a summary in Thurlow, Liu, Albus, and Shyyan (2003), activities in Russian language classes in first and second grades include analyzing the phonological structure of words and syllables; segmenting and blending syllables and phonemes; breaking up sentences into words, words into syllables, and syllables into sounds; determining the order of phonemes in a word; determining the relationship between sounds; determining the number of syllables and phonemes in a word; finding specific phonemes in a word; representing sounds with letters; and putting together

syllables and words. Thus, right from the start Russian children are explicitly taught phonological awareness as well as the basic concepts associated with it.

Activities in reading classes (e.g., Thurlow et al., 2003) include reading aloud, retelling, writing, reciting, and working with written text. Children learn to divide a written text into logical parts, locate essential information in the text, identify the author's communicative goal and message, understand the main idea of the text, explain the means used by the author to communicate his or her ideas, retell the text without omissions, ask and answer questions related to the text, and analyze the text stylistically. Other activities often used in reading classes include making up a story for a set of pictures that has a logical development, talking about one's own experiences analogous to those described in the story, explaining puzzles, and reciting poems. At the end of first grade, children start reading classical literature including the works of Leo Tolstoi, Aleksander Pushkin, and Ivan Bunin.

Besides reading and writing, first-grade children learn many fundamental concepts that underlie language and reading. For example, children learn that language is a means of communication and thought, that reading is a form of communication and a process of understanding written speech involving the author and the reader who act in accordance with their goals, and that there is a difference between literary language and colloquial speech.

By the end of first grade, children are expected to know all letters and sounds as well as the difference between letters and sounds. They are expected to

be able to segment words into sounds, determine the order of phonemes in a word, and understand the difference between vowels and consonants; name soft and hard consonants in words and know how to represent them in writing; know how to use vowels to represent soft consonants; and be able to determine the placement of main stress in a word (Alekseenko et al., 2003). Because smooth and fluent reading is considered the main goal of teaching reading, reading speed is used as the main criterion of achievement in reading assessment. Reading progress is measured three times a year by having a child read a previously unseen passage under timed conditions and then comparing the child's reading rate per minute against reading norms (Kobozeva, Department of Education, Moscow, January 2007, personal communication). By the end of first grade, children are expected to read at a rate of 30 words per minute or higher and demonstrate smooth syllabic reading with elements of whole-word reading with appropriate pauses and intonation (Alekseenko et al.). Recently some methodologists have suggested the use of other criteria of reading achievement in addition to reading speed, including correctness, distinctiveness, and comprehension (Korneyeva, 2000, as cited in Thurlow et al., 2003); however, it is unclear to what extent these criteria have been incorporated into assessment procedures at individual schools.

Russian Research on Literacy

Russian research on literacy has a long history beginning with the pioneering work of Daniil Elkonin in the 1960s (e.g., Elkonin, 1973). A student of

Lev Vygotsky, Elkonin was the first Russian psychologist to observe the crucial role of phonological awareness in children's reading development and to create activities to booster its development (the famous "Elkonin boxes"). Since then Russian psychologists and educators have produced a substantial body of literacy-related research (e.g., Chernyavskaya, 2000; Grushevskaya, 1982; Kornev, 1995; Kreschenko, 2003; Ogarkina, 2002; Velichenkova, 2002; see also Boldyreva & Inshakova, 2004; Inshakova, 2004; Lalaeva, 2004 for a review) and have advanced a number of interesting theories of reading development and dyslexia (see, for example, Tsvetkova, 1997, for a review of theories about the neuropsychological underpinnings of reading and reading disability).

However, Russian research in this area as well as more generally in psychology and education has several characteristics that differentiate it from psychological and educational research in the West. The main differences concern the basic approaches to scientific investigation, the definitions of key concepts, the areas of focus, the methods of assessment, and the overall quality of scholarship. These differences make it difficult to place Russian findings in the context of international research.

General Approaches to Research

Russian scholars working in psychology and related fields have traditionally followed a different approach to scientific inquiry, which is rooted in the work of Luria, Vygotsky, and other prominent Soviet psychologists. The Russian approach

is qualitative in nature, derived from clinical observations in neurology and clinical psychology, and oriented toward single-case study (Glozman, 1999). As such, it stands in stark contrast to the strong psychometric tradition in North American psychology and education, which emphasizes standardized assessment, the use of statistical procedures, and the study of groups (Tupper, 1999).

When applied to the study of literacy and dyslexia, the Russian approach affects both the choice of the target population and the focus of the research. Whereas much of the English-based literature on literacy is based on epidemiological studies with large, representative samples (Snowling, 2000), Russian research in this area has been limited to clinical groups, with the inevitable consequence of sampling bias. The focus of this clinically-oriented work has been on the localization of brain functions and the identification of links in the functional system (Akhutina, 2003) through behavioral observation rather than on the determination of correlates and predictors of literacy achievement through psychometric measurement, which is one of the main directions of literacy research in the West.

More specifically, many Russian studies of literacy development (e.g., Akhutina, 2002; Ogarkina, 2002; Velichenkova, 2002) attempt to connect "certain peculiarities" (i.e., errors, Inshakova, 2004, p. 174) in children's writing, reading, and speech with individual differences in the composition of higher psychological functions and lateral brain organization, which these peculiarities are assumed to reflect (for more on the current approaches in Russia to identifying children with

dyslexia see Inshakova, 2004). This is, for example, how Inshakova describes the causes of writing problems in one group of children with the prevalence of a particular type of speech and spelling errors: “Disturbances of a right hemispheric strategy...prevail when disorders occur with left hemispheric information processing strategy. This allows us to presume that the peculiarities of speech and specific writing and reading disorders of such children are due to right hemispheric differences” (p. 175).

Although this approach presents an interesting alternative to the measurement-based approaches adopted in the West, its usefulness in contemporary literacy research may be limited for several reasons. First, the predominant focus on error analysis in Russian studies appears to be inconsistent with recent findings from both transparent and opaque orthographies showing that many so-called “dyslexic” errors such as letter reversal are a normal feature of reading development (e.g., see Snowling, 2000, for English; Nikolopoulos et al., 2003, for Greek). This inconsistency has also been noted by some Russian psychologists (e.g., Kornev, 2006).

Second, because of the uniquely “qualitative flavor” (Tupper, 1999, p. 4) of Lurian assessment procedures, this approach may be poorly suited for the study of groups, which requires the generation of objective, quantifiable data. Recent attempts to adapt Lurian assessments for use with groups (e.g., Akhutina, 2002) are difficult to evaluate as group characteristics are still typically discussed within a qualitative framework. Third, most neuropsychological assessment methods used to

study literacy development were originally designed for the clinical evaluation of patients with serious brain pathology and it is unclear to what degree these methods may be applicable to the study of normal development. Finally, many Russian assessment methods are based on models of cognitive functioning that were postulated by Russian scientists half a century ago, in an era without MRI and other brain visualization techniques and so far, there have been no attempts to update these models with recent findings from neuroimaging research.

Definitions and Areas of Focus

An important difference between Russian research and the English-based literature concerns terminology. In Russia, reading and writing are studied separately and two different terms are used to refer to problems in each of these areas, the term *dyslexia*, which denotes reading problems, and the term *dysgraphia*, which is used to describe writing difficulties. In the English literature, in contrast, spelling problems are often treated as another symptom of reading disability, and the term dyslexia usually incorporates both reading and spelling difficulties. Further, whereas in the U.S. and many other Western countries dyslexia has been defined both conceptually and operationally, in Russia, there seem to be no clear guidelines for distinguishing between specific reading disability and reading disability in the context of more general learning problems (so-called *reading backwardness*) either in clinical or educational settings. For practical as well as research purposes, all disabled readers are termed dyslexic, regardless of whether

their problems stem from speech, visual, motor, or more global cognitive impairments. Yet the question of pseudo-dyslexia is an important one as children with reading difficulties stemming from more global cognitive deficits (often referred to as *garden-variety* poor readers) may require different diagnostic procedures. This is because performance on many achievement tests including phonological tests is affected, at least in part, by the general ability factor (Frith, 2002) and diagnosing phonological or processing deficits in the presence of low IQ may be difficult. Garden-variety poor readers also appear to respond to reading intervention much better than do truly dyslexic children (Snowling, 2000).

It is also possible for certain disorders to co-occur with dyslexia. For example, Snowling (2000) described two developmental disorders, attention deficit hyperactivity disorder and developmental coordination disorder, which may co-occur with dyslexia, modifying its manifestations. These disorders can cause difficulties in literacy acquisition that are distinct from dyslexia.

The lack of clear definitions and diagnostic criteria for dyslexia is particularly regrettable given the predominance of the neuropsychological perspective on reading difficulties in Russia, a perspective that embraces a medical approach to the study and diagnosis of learning disabilities. On the other hand, the neuropsychological approach itself, with its emphasis on qualitative differences, may preclude the development of an effective definition of dyslexia because there appear to be no qualitative diagnostic symptoms of dyslexia, only quantitative ones, and there may be no qualitative differences between normal readers and readers

with dyslexia (Bruck, 1988; see also Singleton, 2002, Stanovich, 2000, for a review).

Writing development and disorders are accorded much more attention in Russian research than reading development. In fact, most Russian literacy research appears to center on the study of writing and dysgraphia. This is probably because of the widespread belief that writing difficulties are much more common in Russian children than reading ones, a belief that has recently received some empirical confirmation. For example, Inshakova (2004) cites a study of 186 Russian first graders in which the incidence of reading disability was estimated at 10% while the incidence of writing disability was 21%. The greater incidence of writing difficulties over reading ones is perhaps unsurprising given that Russian orthography is transparent for reading but inconsistent for writing (Hamilton, 1980).

Perhaps due to the popularity of clinical models of development as a paradigm for research in developmental and educational psychology, Russian researchers working in these areas appear to be more interested in pathological rather than normal development. This is reflected in the almost exclusive focus of Russian studies on children with learning disabilities and on ways of remediation. In fact, research on normal development represents a relatively new line of inquiry in contemporary psychology and neuropsychology (Glozman, 1999), and attempts to establish benchmarks for normal functioning are just starting to emerge.

Assessment

Traditionally, psychological assessment in Russia has taken two forms: clinical, as a qualitative, process-oriented analysis of—typically—a single clinical case, and criterion-referenced, as a comparison of student achievement against some predetermined—and often arbitrary—cut-point. In literacy development, the former is reserved for assessing abnormal development whereas the latter is used to track students' progress. For example, at the end of first grade, children are expected to demonstrate fluent passage reading at a rate of 30 words per minute (Alekseenko et al., 2003). This is tested by having a child read a previously unseen passage and calculating the child's reading rate. Those consistently failing to reach "the norm" may be referred for clinical assessment to determine the extent and causes of their reading difficulty. The problem with this approach, however, is that the norms appear to be rather arbitrary and are not based on any standardized procedure. As a result, good and poor reading can only be defined in qualitative terms, as reading above or below the norm.

Recently Russian psychologists have started to develop standardized, culture-appropriate tests for evaluating both achievement and cognitive performance, and a number of testing instruments have been published (e.g., Akimova & Kozlova, 2006; Kornev, 1995; Ratanova, 2005). However, most of these instruments have been standardized on small and unrepresentative groups of children from one or two big cities and many lack such important information as reliability and validity data. In discrepancy-based dyslexia assessment, many

psychologists continue to rely on the *Wechsler Intelligence Scale for Children* (Philimonenko & Timofeev, 2001) which was developed in 1949 and modified for use with Russian children in 1973, or more than 35 years ago.

Quality of Scholarship

During the years of communism, Russian psychology like most other areas of science served the needs of political ideology, reflecting through its theories and models the communist belief system. Research articles of that time had a strongly propagandistic tone (Halpern, 1999). Many Soviet scientists intentionally avoided Western sources and concepts, and therefore drifted apart from international research. The collapse of the Soviet Union ended the political isolation. However, 17 years on, the Russian scientific community remains as isolated from world science as it has ever been. Several reasons account for this continued isolation including a lack of financial means, language problems, and a newly emerging anti-Americanism, which has led to a widespread aversion to Western scientific culture (Schiermeier, 2007, p. 527). However, the biggest problems concern factors that are intrinsic to science itself—the overall quality of research, the standards of scholarship, and the principles of specialist training in Russia.

In Russia, there currently seems to be no generally recognized system of research evaluation and no proper peer review ("What scientists say," 2007). Many Russian scientists publish in in-house publications (e.g., university bulletins or working papers published by research organizations) and there is no effective

mechanism to control the quality of what is published. As one Russian physicist has put it in an interview with *Nature*, in Russia, "[t]here is no well defined procedure to distinguish really good science from pure trash" ("What scientists say," p. 529). The notion of peer review is still not accepted by many Russian scientists ("Time," 2007, p. 507), and the idea that scientific reputations should depend on publication lists is often considered incompatible with true science.

The lack of a quality control mechanism in Russian science is not accidental. Russian research institutions are funded by the government through a non-competitive scheme, and the amount of funds allocated to a particular research institute does not appear to depend on the output or performance. Similarly, scientific careers in Russia do not depend on results to the same extent as they do in the West and there is little motivation for Russian scientists to produce research. In fact, according to a database of the scientific productivity of Russian scientists compiled through efforts funded by the Russian Foundation for Basic Research, fewer than one-eighth of all scientists in Russia publish at least one paper a year (Schiermeier, 2007, p. 525). Russia's record of international publications remains particularly bleak as many Russian scientists do not appear to consider international expertise an important attribute of quality science.[6]

---

[6] For example, in their studies of basic word processing in Russian adults reading in a second language, Wade-Wooley (1999) and Abu-Rabia (2001) were unable to cite a single empirical work on Russian word recognition. My own search of several international databases (e.g., ERIC, Academic Search Primer) did not produce any results. Russian researchers also appear to rely predominantly on domestic sources. For example, an examination of 15 recent articles randomly

Low output of high-quality, international-level research may be particularly characteristic of fields such as psychology and education, which have historically been treated in Russia as purely pedagogical disciplines[7] and which have never enjoyed—unlike defense-related fields—generous funding. These disciplines are heavily rooted in Marxist philosophy, materialism, and social-cultural theory and have traditionally emphasized ideology. Until recently, specialist training in these areas was limited to Russian theories of learning and non-quantitative methodologies including literature analysis, experiential reporting, and, in some areas, clinical observation. Western psychological and pedagogical theories as well as preparation in quantitative methods, research design, measurement theory, and statistics were not part of the curriculum. Today the situation is changing and many universities have started offering courses in research methods and statistics as part of their training programs in psychology (Russian Ministry of Education, 2000). However, several years will have to pass before a new generation of psychologists capable of producing systematic, international-level research will emerge in Russia.

At this time, there is a clear dearth of reliable, scholarly research on literacy development in Russia. For example, in a comprehensive review of the literature on early literacy instruction in several different languages compiled by the National

---

selected from a leading Russian journal in psychology, *Issues in Psychology,* revealed that on average, foreign citations accounted for only about 10% of all sources.

[7] Russia did not have a separate PhD degree in psychology until the late 1960s, and psychologists were trained within graduate programs in education.

Center on Educational Outcomes (Thurlow et al., 2003), the authors found only one report of original research conducted in Russia between 1987 and 2000. By comparison, the same review identified 33 empirical studies of reading conducted in China in the same period.

My own extensive search of Russian literature yielded similar results. Using the keywords *reading, writing, disorders, dyslexia, cognitive*, and *comprehension*, I searched three Russian online databases: the database of the Russian Dyslexia Association (RAD) (www.dyslexia.org.ru), the database of a leading Russian journal in psychology, *Issues in Psychology* (www.voppsy.ru), and the database of Pedagogical Library (www.pedlib.ru), an online resource for scholars that provides access to full-text journal articles, books, and dissertations in child psychology and education. Altogether the four databases contained well over 5,000 sources on various topics in education and psychology, covering a 20-year period. Sixty-nine potentially relevant sources were identified; however, only four turned out to be empirical works. A close examination of these studies revealed significant methodological and statistical shortcomings in study design and analysis including comparing the performance of groups using different tests, combining clinical and non-clinical populations in one group, using single tests of ability, and failing to control for confounding variables such as IQ.

Statement of the Problem

In all languages, the ability to read depends on a number of cognitive and linguistic abilities that reflect the processing of spoken words. First and foremost, these include phonological (and especially phoneme) awareness, phonological recoding in lexical access, and verbal short-term memory—abilities that underline the importance of phonology in learning to read. In addition, reading ability (particularly reading comprehension) is influenced by general intelligence and broader language skills such as vocabulary knowledge and listening comprehension.

Yet reading is also a cultural phenomenon as it necessitates the learning of a code of symbols that members of a given culture use to represent speech. This code reflects not only the structural properties of the language for which it developed but also the historical forces that have shaped its development and the influences of other languages on the phonology, morphology, and grammar of the language it represents. As an activity that draws upon an individual's ability to process this culturally determined code, reading is necessarily affected by language-specific factors including the phonological structure of the language, the transparency of the orthography, and the type of information processing required by the particular orthography. These factors appear to influence the ease and speed of reading acquisition such that children learning to read a consistent orthography acquire the alphabetic principle much faster and more easily than do children learning to read an inconsistent orthography. They also affect the choice of strategies readers use in word processing, introducing qualitative differences into reading acquisition across

languages. For example, individuals reading more consistent orthographies have been shown to rely almost exclusively on serial, letter-by-letter decoding, whereas persons reading inconsistent orthographies are forced to supplement decoding with processing at the level of syllables, rhymes, and whole words. Furthermore, language-specific factors appear to influence the manifestation of reading disability as well as the degree of impairment that such a disability imposes on the sufferer, making some languages more "dyslexia-friendly" (Frith, 2002) than others.

The role of language-specific factors in the development of the reading system has several important implications. First, it underscores the importance of developing language-specific models of reading and reading disability that can accommodate the specific features of a given language. Second, language-specific factors make it inappropriate to disentangle research findings from the specific language context in which they were obtained, rendering such findings largely non-generalizable across languages. Third, the extent to which specific cognitive abilities can be useful in predicting reading achievement appears to depend on the features of the particular language, especially its orthography. These features may also affect the underlying structure of some of these cognitive abilities (e.g., phonological awareness) as well as the usefulness of the specific tasks used to measure them.

Literacy research in the Russian language has a long history. However, Russian research in this area has been distinct from international research in its approaches to scientific investigation, the areas of focus, and the terminology and

methods of assessment used. These differences make it difficult to place Russian findings in the context of international research on literacy. Russian research also appears to be limited to the study of abnormal development and literacy disorders, particularly dysgraphia, while the normal course of reading development remains largely unexplored. More specifically, there have been no attempts to determine, through systematic statistical investigation, the relative importance of various cognitive and linguistic factors to reading acquisition in the Russian language as well as to clarify to what extent Russian readers are similar to those in other languages. This knowledge is crucial to the construction of predictive models of reading in the Russian language and the development of appropriate assessment tools.

The purpose of this study is to explore a specific set of concurrent cognitive predictors of reading ability in a group of Russian-speaking beginning readers. Specifically, I examined the relative influence of phonological awareness, rapid naming, verbal short-term memory, and IQ in the prediction of decoding accuracy and decoding rate in Russian in order to identify the strongest independent predictors and to determine to what extent these predictors are similar to those that have been identified in other languages. A second purpose is to examine the structure of phonological awareness in the Russian language in order to clarify the hierarchy of difficulty of linguistically and cognitively diverse tasks and determine whether they form a single variable. A third purpose of this study is to test the power of the Simple View of reading in Russian by examining whether linguistic

comprehension abilities (i.e., listening comprehension and vocabulary) significantly add to the prediction of reading comprehension after differences among children in IQ have been statistically eliminated and whether decoding skill contributes any significant variance to the prediction over and above that contributed by IQ and linguistic measures.

A fourth purpose of this study is to examine the cognitive and linguistic skills of the top and bottom readers in the sample in order to determine the largest sources of difference. For this purpose, the sample will be split into two groups using decoding rate as a criterion, and the groups will be compared on IQ, phonological awareness, listening comprehension, vocabulary, and reading comprehension. The children will then be classified as good and poor readers based on their scores on the five predictors. The final purpose of this study is to examine the cognitive, linguistic, and reading profile of the poorest readers in the sample in order to identify their relative strengths and weaknesses.

The following research questions have been formulated.

1. a. What is the size of the overall relationship between decoding accuracy and the set of phonological awareness, verbal short-term memory, rapid automatized naming, and IQ in beginning Russian readers? b. How much of the relationship is contributed uniquely by each predictor?
2. a. What is the size of the overall relationship between decoding rate and the set of phonological awareness, verbal short-term memory, rapid

automatized naming rate, and IQ in beginning Russian readers? b. How much of the relationship is contributed uniquely by each predictor?

3. Do rhyme and phoneme detection, phoneme segmentation, and phoneme deletion tasks tap the same underlying ability in the Russian language?
4. How accurately can reading comprehension be predicted from (a) IQ, (b) listening comprehension and vocabulary over and above IQ, (c) decoding rate over and above IQ, listening comprehension, and vocabulary, and (d) the combination of IQ, listening comprehension, vocabulary, and decoding rate?
5. a. To what degree are there differences between good and poor readers in the population on linear combinations of IQ, phonological awareness, listening comprehension, vocabulary, and reading comprehension? b. Can the children in the sample be correctly classified as good and poor readers based on their scores on the five predictor variables?
6. What is the cognitive, linguistic, and reading profile of the poorest readers in the sample?

CHAPTER 3

METHODS

Participants

Seventy-nine first- and second-grade Russian children participated in the study. There were 26 first graders and 53 second graders. The two groups of children were comparable in age, socio-economic status, and reading level because they were separated by only a three-month summer break. City-wide reading norms for both exiting first graders and entering second graders are equivalent, and public schools in Russia do not typically offer summer school programs, so most children do not take any classes in the summer. All the participants spoke Russian as their native tongue.

The children came from four intact classes[8] in three public elementary schools, each in a different school district in Moscow. All the first graders (n = 26) and some of the second graders (n = 16) came from the same school; the rest of the second graders came from two different schools (n = 12 and n = 25, respectively). There were 47 boys (15 first graders and 32 second graders) and 32 girls (11 first graders and 21 second graders). There were more boys than girls in the sample because many girls in the classes tested turned out to be non-native Russian speakers and were therefore excluded from the study. All the children had normal

[8]In each class, 10-30% of the children were non-native speakers of Russian and/or did not speak Russian at home. These children were excluded from the study.

hearing; none appeared to have any neurological or emotional problems. The children ranged in age from 80 to 117 months; the average age of the first graders was 92 months and that of the second graders was 96 months. Such a range is not uncommon in Russian schools as Russian parents have a choice of sending their children to school at the age of six or seven.

The schools were selected because they were deemed to be ordinary Moscow public schools and because I had access to them. All three used the same academic curriculum and taught the same subjects, with an equal number of hours dedicated to reading and writing instruction and with an equal number of classes per day held between 8:30 and 12:00. Reading instruction was uniform across the participating schools, and all three used the same textbooks in Russian language and Reading classes. The schools represented a lower-middle to upper-middle socioeconomic class population. This was determined in consultation with the principals of the schools based on the children's records.

The goals, design, and procedure of the study were explained to the principals of the participating schools and the homeroom teachers of the classes tested and their consent to conduct the study was obtained. The teachers then informed the parents and obtained their verbal consent. In Russia, it is unusual for a parent to refuse participation in testing organized or consented to by the school; there were no refusals to participate.

Instrumentation

The testing battery consisted of 15 tests: five phonological tests, a nonword repetition test, a forward digit span test, two RAN tests, a word reading test, a nonword reading test, a vocabulary test, a listening comprehension test, a reading comprehension test, and a test of nonverbal ability (see Appendix P for the original battery in Russian). All tests except vocabulary and nonverbal ability were specifically designed for this study; commercially available tests were used to measure vocabulary and nonverbal ability. The tests are described below.

Phonological Awareness

Five phonological tests were used in this study: Four assessed different aspects of phonemic awareness and one assessed rhyme awareness (see Appendices A-E). The tests were based on Adams's (1990) hierarchy of task difficulty and were modeled after those commonly described in the reading acquisition literature (e.g., see de Jong & van der Leij, 2003; Gottardo, 2002; Muter & Snowling, 1998; Yopp, 1988, for some common tests of phonological awareness). The tests were designed based on feature analysis of the sounds constituting test items so that all commonly occurring places and manners of articulation of Russian consonants and all locations of Russian vowels were represented (Hamilton, 1980). Wherever real words were used, they were selected from first- and second-grade Russian language textbooks (Kubasova, 2003, 2005; Zelenina & Khokhlova, 2004) to ensure their familiarity to the children. Prior to testing, the tests were piloted on a small group

of Russian-speaking children (three native speakers and three non-native speakers) to ensure clarity of instructions and word familiarity. The stimuli were presented orally by the same examiner to all the children. Each test was preceded by one example and one or two practice trials. The items in each test were scored as either correct or incorrect. The five tests were combined to create a composite measure of phonological awareness, and interval-level ability measures were derived for each child using Rasch analysis (see Analyses).

Rhyme Detection

The rhyme detection test (see Appendix A) tapped the children's sensitivity to rhyme. The children were presented with a set of three words, two of which rhymed and one of which did not, and asked to identify the non-rhyming word. The test consisted of 18 sets of high-frequency monosyllabic words with the following structures: CVC (six sets), CCVC (six sets), and CVCC (six sets). The rhyming words differed from each other in the initial phoneme only; the non-rhyming word differed from the rhyming ones in the initial phoneme and either in the vowel (nine items) or in the final consonant (nine items). The non-rhyming word was placed randomly in the first, second, or third position in each set. The children received the following instructions: "Listen as I say three words. Two of them sound similar, but the third one does not. Tell me which one sounds different from the other two."

Initial Oddity

The initial oddity test (see Appendix B) measured the children's ability to isolate initial sounds in spoken words. The children listened to three words in a set and selected the word that began with a different sound from the other two. All words were common nouns and all began with a consonant. There were five sets of CVC words and five sets of CCVC words. The words in each set did not rhyme, and the CCVC words began with a different cluster. The "odd" word was placed randomly in the first, second, or third position. The children received the following instructions: "Listen as I say three words. Two of them start with the same sound, but the third one starts with a different one. Tell me which one starts with a different sound."

Final Oddity

The final oddity test (see Appendix C) measured the children's ability to isolate final sounds in spoken words. The children listened to three words in a set and selected the word that ended with a different sound from the other two. All words were common nouns and all began with a consonant. There were five sets of CVC words and five sets of CVCC words. The words in each set differed in the initial phoneme and the vowel; the CVCC words also differed in the final cluster. The "odd" word was placed randomly in the first, second, or third position in each set. The children received the following instructions: "Listen as I say three words. Two

of them end with the same sound, but the third one ends with a different one. Tell me which word ends with a different sound."

Phoneme Segmentation

A tapping task was used to assess phonemic segmentation skills (see Appendix D). The child was given a blunt pencil and asked to segment a spoken word into its component phonemes by tapping out the number of phonemes in the word. The examiner counted the taps. The test had 13 items: three each of CVC, CCVC, CCVCC, and CCCVC items, and one CCCVCC item. All were monosyllabic nonwords that conformed to Russian pronunciation rules. The nonwords were used to minimize the use of orthographic knowledge to assist phonemic segmentation (Nation & Hulme, 1997). The following instructions were given for the test: "I am going to say some made-up words. Listen and when I finish saying a word, tap out the number of sounds in it."

Phoneme Deletion

In the phoneme deletion test (see Appendix E), the children were presented with nonwords and were asked to take away a specified sound from each nonword and to say the new word.[9] The resulting words were all real, high-frequency words commonly found in first- and second-grade textbooks and likely to be familiar to

[9] In English, such tasks are typically presented in the following format: "What word would be left if /f/ were taken away from *fice*?" The correct response is *ice*.

Russian-speaking children. Nonwords were used to control for possible effects of lexical familiarity (Gottardo, 2002). The test included 12 monosyllabic nonwords, four of which required the deletion of an initial phoneme (two CVC and two CCVC items), four the deletion of a final phoneme (two CCVC and two CVCC items), and four the deletion of a phoneme in the middle (two CVC and two CVCC items). Four nonwords required the deletion of a consonant followed by a vowel and eight required the deletion of a consonant from a consonant cluster. The following instructions were given for the test: "I'm going to say some made-up words and I want you to drop a particular sound from each word. I'll tell you which one. And then I want you to say the new word. For example, if I say "шлось" [shlos'] and ask you to drop the "ш" [sh], what will be left?"

Eliminated Tests

Originally, the phonological battery included two more tests that are typically used in studies of phonological awareness, a test of initial phoneme identification (IPI), in which a child is presented with a picture of a familiar object and is asked to say the first sound of the object's name, and a test of phoneme blending (PB), in which a child is asked to blend separately presented sounds into words. These tests were eliminated after piloting: The IPI test was confusing to the children, who tended to name letters rather than sounds, and the PB test was too easy, with ceiling effects even for the non-native speakers.

Verbal Short-Term Memory

Two tests were used to assess the short-term retention of speech-based information, a nonword repetition test and a forward digit span test (see Appendices F and G). The tests were preceded by an example and a practice trial. The items in each test were scored as either correct or incorrect. The two tests were combined to create a composite measure of verbal short-term memory, and interval-level ability measures were derived for each child using Rasch analysis (see Analyses).

Nonword Repetition

The nonword repetition test (see Appendix F) was modeled after those described in Gottardo (2002) and Wade-Wooley (1999). The children were asked to repeat 24 nonwords with progressively more difficult sound structures. The nonwords began at one syllable and ended at four syllables and were presented in order of increasing difficulty. To obtain a better measure of short-term retention of verbal information, the nonwords were constructed based on English nonwords (Muter & Snowling, 1998) so as to minimize their likeness to real words and were modified to accommodate Russian pronunciation. There were six each of one-, two-, three-, and four-syllable nonwords. The stimuli were presented in a clear, monotone voice and were not repeated. The children received the following instructions: “I’m going to say some made-up words. Listen carefully and after I finish saying a word, repeat it as best as you can.”

Forward Digit Span

The forward digit span test (see Appendix G) was modeled after Turner's Digit Memory Test (Turner & Ridsdale, 2004). The children were asked to repeat digit sequences that increased from 2 to 9 digits. In Russian, digits 2, 3, 5, 6, and 7 are monosyllabic; digits 1, 8, and 9 are bi-syllabic; and digit 4 is tri-syllabic. All the digits were used in the test to avoid the repetition of digits within longer sequences. The test had two trials at each digit sequence length. Digits were presented orally in an even monotone at the rate of one digit per second. The test was discontinued after a child had failed both trials of a digit sequence. The following instructions were given: "Listen carefully as I say some numbers. They will get longer and longer. Repeat each number as best as you can."

Rapid Automatized Naming

Two RAN tests were used to assess the speed of processing required to retrieve lexical items: object naming and digit naming. The tests were modeled after those described in Wolf, Bowers, and Biddle (2000). The scores on the two measures were combined to create a composite measure of rapid naming. The composite measure was also transformed into an items-per-minute measure (RAN rate) for comparison with decoding rate.

RAN-Object

In the RAN-object test (see Appendix H), the children were shown color pictures of five common objects and asked to name them as fast as possible. The five objects were a chair, a bone, a ball, an umbrella, and a cake. The pictures were arranged in random order in a 10 x 5 array on a single page. All words were monosyllabic and all began with a different consonant. One had a CCVC structure, one had a CVC structure, and three had a CVCC structure. Before the administration of the test, two practice trials were given. The first trial introduced the five objects and the child was asked to name them. In the event of difficulty, I provided the target word. The second trial consisted of two lines with the five objects each and introduced the child to the rapid naming procedure. After the second trial, the child proceeded to complete the task. The time used to name the stimuli was measured with a stopwatch. The test was scored as the total naming time in seconds.

RAN-Digit

The RAN-digit test (see Appendix I) was identical to the RAN-Object test except that digits were used instead of pictures of objects. The following five digits were used: 2, 3, 5, 6, and 7. These digits were selected because they are represented by monosyllabic words in Russian. The instructions and administration procedure were the same as in the object-naming test.

Decoding Ability

Two tests were constructed to measure decoding ability, a word reading test and a nonword reading test. The stimuli for these tests came from six Russian textbooks approved and recommended for use in public schools by the Ministry of Education of the Russian Federation. The textbooks were *Literature Reading, Grade 1* (Kubasova, 2003), *Literature Reading, Grade 2, Part 1* (Kubasova, 2005), *Literature Reading, Grade 3, Part 1* (Kubasova, 2004), *Russian Language, Grade 1* (Zelenina & Khokhlova, 2004), *Russian Language, Grade 2, Part 1* (Ramzaeva, 2005a), and *Russian Language, Grade 2, Part 2* (Ramzaeva, 2005b).

Word Reading

The word reading test (see Appendix J) measured word decoding ability. Four word lists were prepared with a total of 50 words. Single words were used rather than a passage to eliminate context effects in order to obtain a purer measure of decoding ability (Aaron & Joshi, 1992). The words were randomly selected from a larger pool of 472 words compiled by sampling several *Russian Language* and *Literature Reading* textbooks for Grades 1-3 (see above). All the words were high-frequency and were likely to be familiar to first- and second-grade children. List 1 contained 20 words; there were five each of one-, two-, three-, and four-syllable words. The one- and two-syllable words were morphologically simple, consisting of a root or a root and an ending. The three-and four-syllable words were morphologically complex and contained derivational morphemes. Lists 2 and 3

were made up of words taken from a special section in the *Literature Reading* textbooks for Grades 2 and 3 that teaches children to recognize orthographically complex Russian words by sight, using a whole-word strategy. All words in this section are presented in pairs—a morphologically simple word (i.e., consisting of just a root) and a morphologically complex word derived from the simple one by adding various suffixes and/or prefixes. List 2 contained ten morphologically simple words—five consisting of one syllable and five consisting of two syllables, and List 3 contained ten morphologically complex words—five consisting of three syllables and five consisting of four syllables—all of which were derivatives of the words on List 2. List 4 contained ten morphologically complex words taken from the Syllable Reading section in the *Literature Reading* textbooks. This section teaches children to decode difficult Russian words syllable by syllable. There were five three-syllable and five four-syllable words on this list.

Four stimulus cards were prepared, one for each word list. The words were type-written in 20-point font, double-spaced, numbered, and arranged in rows, with two words per row. The children were given the cards one at a time, and were asked to read the words on each card as quickly as they could and without errors.

Two measures, reading speed and reading accuracy, were taken separately for each list. Reading speed was measured with a stopwatch; reading accuracy was marked by assigning a score of 1 for a correctly read word and a score of 0 for an incorrectly read word. Errors were defined as repetition, substitution, deletion, or addition of one or more sounds or syllables in a word; reversal of letters or

syllables in a word; and semantic substitution (e.g., reading *office* for *official*). The placement of the main stress on the wrong syllable was also considered an error. Self-corrections were accepted as correct answers.

Nonword Reading

The nonword reading test (see Appendix K) measured nonword decoding ability. In the developmental literature, it is considered a more basic measure of decoding than word reading tests (Snowling, 2000). Two lists of nonwords were prepared. List 1 contained 24 nonwords that were visually dissimilar to real words. There were six one-syllable words created by changing the last consonant in real one-syllable words; six two-syllable words created by transposing the syllables in real two-syllable words with CV and CVC syllable structures; six three-syllable words created by transposing the last two syllables in real three-syllable words; and six four-syllable words created by slightly modifying English nonwords. The real words on which the one-, two-, and three-syllable nonwords were based were taken from three textbooks, Grade 1 *Russian Language* (Zelenina & Khokhlova, 2004), Grade 1 *Literature Reading* (Kubasova, 2003), and Grade 2 *Literature Reading* (Kubasova, 2005). The four-syllable nonwords were taken from the nonword repetition test in Muter and Snowling (1998) and modified to accommodate Russian pronunciation. Prior to inclusion in the list, the nonwords were shown to a small group of native Russian speakers to verify their pronounceability, conformity to Russian spelling patters, and dissimilarity to real words.

List 2 contained 24 nonwords that were visually similar to real words. These nonwords were created by changing the stressed vowel in common Russian words taken from the *Russian Language and Literature Reading* textbooks for Grades 1 and 2. There were six each of one-, two-, three-, and four-syllable nonwords. Prior to inclusion in the list, the nonwords were shown to a small group of native Russian speakers to verify their pronounceability, conformity to Russian spelling patters, and similarity to real words.

Two stimulus cards were prepared, one for each list. The words were type-written in 20-point font, double-spaced, numbered, and arranged in rows, with two words per row. The words were presented as made-up words. The children were given the cards one at a time and were asked to read the words on each card as quickly as they could and without errors.

Two measures were taken separately for each list, reading speed and reading accuracy. Reading speed was measured with a stopwatch; reading accuracy was marked by assigning a score of 1 for a correctly read item and a score of 0 for an incorrectly read item. Errors were defined as repetition, substitution, deletion, or addition of one or more sounds or syllables in a word; reversal of letters or syllables in a word; and substitution of a target nonword with a real word. Self-corrections were accepted as correct answers.

Combined Measure of Decoding

Previous researchers (e.g., Nation & Snowling, 2004; Stanovich et al., 1984) have sometimes treated word and nonword reading as measures of different abilities (i.e., word recognition vs. grapheme-phoneme decoding) and even used nonword reading to predict word reading. In this study, however, the two measures were hypothesized to reflect one underlying ability because the children were beginning readers, who were unlikely to be visually familiar with many of the more complex items on the word reading test. Because Russian orthography is transparent for reading, the reading mechanism for these unfamiliar words should be identical to that for nonwords, i.e., decoding. In this sense, the difference between reading words and reading nonwords boils down to that between reading visually familiar and visually unfamiliar words. Perhaps for this reason, combining word and nonword reading has been an acceptable strategy in languages with transparent orthographies such as Dutch (de Jong & van der Leij, 2002) or German (Wagner et al., 1997).

To verify whether the two tests did indeed measure the same construct, Rasch analyses were conducted with the word reading test and the word reading and nonword reading tests combined and intercorrelations were computed between word reading rate, nonword reading rate, and the combined decoding (i.e., word and nonword reading) rate. The results are reported in the Analyses Chapter.

Reading Comprehension

The reading comprehension test (see Appendix L for the original test and Appendix N for the English translation) was designed using the following procedure. First, all short (i.e., half a page or less) stories from Grade 1 and Grade 2 *Reading* and *Russian Language* textbooks (Kubasova, 2003, 2005; Ramzaeva, 2005a, 2005b; Zelenina & Khokhlova, 2004) were entered into a computer file. This resulted in 28 stories. The stories were then analyzed using three readability indices: number of words per text, grade level (i.e., GL, the number of years of formal education needed to easily understand the text on first reading), and reading ease (RE). Five short stories were selected; they were written at a grade level of 1.0-2.6 and had a high reading ease score: text 1, 27 words, GL 1.0, RE 100; text 2, 39 words, GL 1.4, RE 100; text 3, 32 words, GL 1.7, RE 100; text 4, 54 words, GL 1.8., RE 98.6; text 5, 29 words, GL 2.6, RE 98.4.

The stories were arranged in the order of difficulty from the simplest to the most difficult using grade level as a criterion. They were titled Story 1, Story 2, and so on to avoid giving any hints about the content. Each story was printed in a 14-point font on an A-4 sheet of paper. For each story, 2-4 questions covering both main ideas and additional information were written. In all, there were 16 questions; eight required literal comprehension and recall of information and eight required making interpretations or drawing inferences by analyzing relationships, reflecting on the overall meaning of the story, or using world knowledge.

The stories were presented on a separate sheet of paper and the children were asked to read them aloud. If a child made a mistake (e.g., mispronounced or omitted a word), it was corrected immediately by the examiner. Each story was followed by the questions, which were presented orally. The children were not allowed to see the story when answering the questions. All answers were tape-recorded and later transcribed verbatim. The transcripts were double-checked by a research assistant for correct entry. A holistic scoring scheme was developed (see Table 1), and answers were scored independently by two raters. The inter-rater reliability was 98%. The scores were then transformed into Rasch person ability logits (see the Analyses Chapter).

Table 1. *Reading and Listening Comprehension Scoring Guide*

Score	Criteria
2	Response is accurate and complete showing adequate understanding of the story. Inferences about characters and/or events are logical and are based on the information in the story. Sentences are well formed and vocabulary is appropriate. Response is based solely on the information in the story.
1	Response is accurate but incomplete: It shows an overall understanding of the story but omits some information. Inferences about characters and/or events are generally logical but may be far-fetched. Sentences are generally well formed but may contain grammar and vocabulary errors. Response may include information not mentioned in the story.
0	Response is irrelevant or wrong or there is no response even after several prompts. Response shows little or no understanding of the story and/or includes mostly irrelevant information. Inferences are irrelevant or wrong.

Linguistic Abilities

Listening Comprehension

The listening comprehension test (see Appendix M for the original test and Appendix O for the English translation) was developed using the same text pool and the same procedure as those used in the construction of the Reading Comprehension (RC) test. The five texts selected for this test were similar to those selected for the RC test in all three readability statistics: text 1, 33 words, GL 1.0, RE 100; text 2, 28 words, GL 1.4, RE 100; text 3, 27 words, GL 1.7, RE 100; text 4, 78 words, GL 1.5, RE 100; text 5, 29 words, GL 2.6, RE 98.4. Each story was followed by 2-4 questions covering main ideas and additional information. There were 16 questions; eight required literal comprehension and recall of information including details and eight required making interpretations or drawing inferences by analyzing relationships, reflecting on the overall meaning of the story, or using world knowledge.

The stories were presented orally, and the children were instructed to listen and to answer the questions. Answers were tape-recorded and later transcribed verbatim; the transcripts were double-checked for correct entry by a research assistant. The test was scored using the same scoring scheme as the one developed for the RC test (see Table 1). A preliminary analysis of the test revealed that one item (Question 13) did not lend itself to partial credit scoring. This item asked for a specific person's name and could only be scored either "correct" or "incorrect." This item was deleted from the test. Answers to the reaming 15 questions were

scored independently by two raters. The inter-rater reliability was 98%. The scores were then transformed into Rasch person ability logits (see the Analyses Chapter).

Vocabulary Knowledge

Expressive vocabulary was assessed using the Vocabulary Subtest of the *Wechsler Intelligence Scale for Children* (WISC), Russian version (Philimonenko & Timofeev, 2001). The test consisted of 40 progressively more difficult vocabulary items that the child was asked to define. There were 33 nouns, three verbs, and four adjectives on the test. Although the testing manual includes norms for Russian children, a decision was made to use raw scores rather than the provided norms for the following reasons.

First, according to the manual of the Russian version (Philimonenko & Timofeev, 2001), the original test was translated and adapted for use in Russia in 1973, or more than 35 years ago. Since then the test has not been re-evaluated or re-normed. Second, it appears that the test has never been formally standardized on a representative group of Russian children and that the U.S. norms are used instead. The use of these norms is explained in the manual on the grounds of a 1973 study by Panasyuk, who assessed the reliability of the Russian version of the WISC by administering it to a group of mentally retarded children aged 6-13 ($n = 67$) and a group of controls of the same age. The obtained results, according to the manual, showed that "the scores of the normal children were entirely within the norms suggested by D. Wechsler [and that] the results obtained for the normal and for the

retarded children differed statistically significantly ($p < 0.001$) on both the separate subtests and on the full test" (p. 10). The authors of the Russian manual concluded that, "the equivalence of the original test and the adapted version of the WISC...can thus be considered established and the use of D. Wechsler's norms justified" (p. 11). However, such an equating procedure does not constitute a substitute for standardization and can hardly justify the use of the U.S. norms for Russian children. Finally, although not stated explicitly in the manual, it appears that the U.S. norms used in the Russian version are those obtained during the *original* standardization in 1949.

At the beginning of the test, the children were given the following instructions: "I'm going to say some words and I want you to tell me what each word means," which was followed by the question "Could you explain what [a bicycle] is?" This question was used for all test items. If a child did not answer within about 10 seconds, the question was repeated and, in case of no response, followed by additional questions (e.g., What can you say about…? Can you describe…?). All answers were tape-recorded for verification. Testing was discontinued after four consecutive failures.

The test was scored as described in the Russian manual (pp. 30-31), using suggested responses as a guide. First, each response was assigned either a (+) (i.e., the target word is familiar to the child) or a (-) (i.e., the target word is unfamiliar to the child). Answers assessed as (-) were awarded 0 points. Answers assessed as (+) were awarded either 1 or 2 points using the following criteria: Answers giving a

precise definition of the target word or a close synonym were awarded 2 points, whereas answers giving a correct but incomplete definition, a description of certain features of the target word, or an example demonstrating understanding of the target word without an accurate definition were awarded 1 point. This way, partial credit could be awarded for partly accurate or accurate but incomplete answers. Several words on the test had two meanings. In these cases, all meanings were accepted as correct regardless of the suggested answers in the manual. Two raters scored the test. The inter-rater reliability was 97%. Disagreements were resolved through discussion and by comparing the definition in question against that given in a Russian language dictionary (Ozhegov & Shvedova, 2003). The raw scores were transformed into Rasch person ability logits (see the Analyses Chapter).

Nonverbal Ability

The Matrices Subtest of the *Kaufman Brief Intelligence Test*, 2nd ed. (KBIT-2, Kaufman & Kaufman, 2004) was used as a measure of nonverbal ability. According to the manual, the Matrices subtest measures a person's ability to use nonverbal reasoning to solve novel problems. This 46-item nonverbal measure is composed of several types of visual stimuli, some of which are concrete (e.g., people, flowers) and others abstract (e.g., symbols). All items require an understanding of the relationships among the stimuli, and all are multiple-choice, requiring the examinee either to point to the correct response or to say its letter. To minimize chance guessing, each item includes at least five response options and

many include six. The test consists of three sections, each preceded by a trial item. For the easiest items, the child selects one of five pictures that goes best with a stimulus picture (e.g., a car goes with a truck). For the next set of items, the child chooses one of six pictures that best completes a 2 x 2 visual analogy (e.g., a carrot goes with a rabbit). The remaining items involve abstract stimuli and require the child to solve a 2 x 2 or 3 x 3 matrix. Many items also require that the child keep several relationships in mind simultaneously. According to the test manual, the Matrices Subtest is appropriate for non-English speaking populations and its directions may be presented in a foreign language.

The test was administered in accordance with the guidelines described in the manual. Children were instructed to point to the picture they thought represented the correct answer. Testing was discontinued after four consecutive failures. Raw scores were transformed into Rasch person ability logits (see the Analyses Chapter).

Procedures

The second graders were tested in the first two weeks of September, right after the beginning of the school year; the first graders were tested in the last two weeks of May, right before the end of the same school year. This was done to ensure the comparability of the children. The children were tested at their respective schools in quiet rooms during school hours. All children were tested individually in two sessions, each lasting approximately 45 minutes. Most of the

children were tested in two consecutive sessions on the same day with a 20-minute break between the sessions. A small number of children were tested over two days because they had to leave early and could not finish all the tests. The children were rewarded for participation with pencils.

Every effort was made to ensure that the testing sessions were not interrupted and that the children were not distracted. Unfortunately, avoiding distractions and interruptions was not always possible. For example, testing was occasionally interrupted by the bell ringing, or by teachers coming in and asking questions. Also, some children had to be tested during a recess, when the noise level was higher than during class time. In these cases, if the examiner felt that the child may have been distracted by something in the environment, the question was repeated. If a child was interrupted while performing a speeded task, the task was re-administered.

I and a research assistant administered the tests. A different research assistant was used at each school, but all were qualified professionals with some training in psychology. The children were sent to the testing room by their homeroom teacher, one at a time. Administration of the test battery was preceded by questions about the child's favorite subject and friends to put the child at ease. I then told the child that I was at the school to learn what elementary school children knew about sounds and how well they could read and that over the next two hours, we would play word games, read, listen to stories, and look at pictures. I asked the child to let me know if he/she wanted to take a break. An overwhelming majority

of the children were very enthusiastic about the testing and appeared happy to participate. Many asked questions about the testing and made comments about the tests. A few children appeared to be overly shy; in these cases, an effort was made to establish rapport with the child. There were no uncooperative children.

Table 2. *The Testing Battery*

Block	Test and number of items	Scoring
Phonological Awareness	Rhyme detection, 18	Rasch ability logits
	Initial oddity, 10	Rasch ability logits
	Final oddity, 10	Rasch ability logits
	Phoneme segmentation, 13	Rasch ability logits
	Phoneme deletion, 12	Rasch ability logits
Short-Term Memory	Nonword repetition, 24	Rasch ability logits
	Forward digit span, 16	Rasch ability logits
RAN	RAN-object, 50	Naming speed
	RAN-digit, 50	Naming speed
Word Reading	Word lists 1-4, 50	Speed & Rasch ability logits
Nonword Reading	Nonword lists 1 & 2, 48	Speed & Rasch ability logits
Reading Comprehension	Reading Comprehension, 16	Rasch ability logits
Listening Comprehension	Listening Comprehension, 16	Rasch ability logits
Vocabulary	Vocabulary Subtest, WISC, Russian version, 40	Rasch ability logits
Nonverbal Ability	Matrices Subtest, KBIT-2, 46	Rasch ability logits

Note. KBIT-2 = *Kaufman Brief Intelligence Test*, 2nd ed. (Kaufman & Kaufman, 2004); WISC = *Wechsler Intelligence Scale for Children* (Philimonenko & Timofeev, 2001).

Prior to testing, all tests were divided into nine blocks (Table 2). The Phonological Awareness, Short-Term Memory, RAN, Word Reading, Nonword

Reading, and Listening Comprehension blocks were administered by me; the Reading Comprehension, Vocabulary, and IQ blocks were administered by one of the research assistants. Thus, tests that required that participants analyzed orally presented stimuli were always administered by the same person and under the same voice conditions.

The blocks were administered in the same fixed order to all the children. The Phonological Awareness block was administered first, followed by the Short-Term Memory, RAN, Word Reading, Nonword Reading, Listening Comprehension, IQ, Vocabulary, and Reading Comprehension blocks. This order made most sense from the point of view of sustaining the child's interest and motivation because it enabled the examiners to alternate tests requiring a lot of cognitive effort and concentration (e.g., phonological awareness tests) with those that were seemingly easy and fun (e.g., RAN and IQ tests). The tests in the Phonological Awareness block were administered in the following order, which was hypothesized to reflect test difficulty from easier to more difficult: rhyme detection, initial oddity, final oddity, phoneme segmentation, and phoneme deletion. In the Short-Term Memory block, the nonword repetition test was administered first, followed by the digit span test. In the RAN block, RAN-object was administered first, followed by RAN-digit. The order of the items on the phonological tests was randomized; the items on all the other tests were administered in a fixed order that reflected a hypothesized continuum of difficulty.

The IQ and vocabulary tests were discontinued after four consecutive failures; all the other tests were administered in full regardless of the child's performance.

Some of the tests appeared to be rather difficult conceptually for some of the children; these children required more detailed explanations and demonstration. The items on the phonological awareness tests as well as the questions on the reading comprehension, listening comprehension, and vocabulary tests were repeated as needed. The items on the memory tests were never repeated.

CHAPTER 4

ANALYSES

Analyses were conducted in five steps. First, the reading tests were examined separately and as a combined measure to determine whether they tapped the same underlying ability and could therefore be used as a single measure. Second, seven variables were constructed using Rasch analysis for the cognitive, linguistic, and decoding abilities for which dichotomous or partial credit data were collected. These variables were Phonological Awareness, Short-Term Memory, Decoding Accuracy, Listening Comprehension, Reading Comprehension, Vocabulary, and Nonverbal Ability. Rasch analysis was also used in this step to answer Research Question 3. Third, the assumptions of multivariate analyses were assessed through SPSS FREQUENCIES and SPSS DISCRIMINANT and necessary adjustments were made (see pp. 173-176). Multiple regression analyses were then performed to answer Research Questions 1, 2, and 4, and direct discriminant function analyses were performed to answer Research Question 5. Finally, the children's *z*-scores on the reading measures were examined to identify the poorest readers in the sample in order to investigate their cognitive and linguistic abilities.

Decoding Measures

To verify that the reading tests used in this study measured the same construct and could be combined, Rasch analyses were conducted with the word

reading test and the word reading and nonword reading tests combined. Accuracy data were used in these analyses. The results of the principal components analysis of standardized residuals for word reading and for the combined measure were similar: The amount of variance explained by the Rasch measure was 50.4% for word reading and 50.9% for the combined measure. Item reliabilities were also comparable, .81 for word reading and .84 for the combined measure; person reliability was significantly higher for the combined measure (.86 vs. .60). There were slightly more misfitting items on the combined measure (three vs. one on word reading using the .5-1.7 range for acceptable fit), but the error estimates were lower for the combined measure than for word reading. Therefore, the combined measure was used in the main analyses.

The reading speed and accuracy data were also converted into three items-per-minute measures, word reading rate (WRATE), nonword reading rate (NRATE), and combined decoding rate (DRATE), by dividing the number of items read correctly by the total reading speed on the respective test and multiplying the result by 60. The Pearson correlations between DRATE and NRATE and DRATE and WRATE were virtually identical, .94 and .93, respectively. The correlation between WRATE and NRATE was somewhat lower, but still acceptable at $r = .76$. Therefore, the combined measure of decoding rate was used in the main analyses.

Latent Variable Construction: Rasch Analysis

The Rasch model (one-parameter logistic model) is a mathematical probability model that provides a practical way to construct linear item calibrations and person measures from ordinal observations (Wright & Stone, 1979) and to investigate the dimensionality of the constructed variables (Smith, 1996). The Rasch model is based on two assumptions: (a) that a more able person has a higher probability of success on any item, and (b) that any person has a higher probability of success on an easy item than on a difficult one (Lunz, Stahl, & Wright, 1996). A key feature of the Rasch measurement model is the requirement of unidimensionality, or that a testing instrument measure only one underlying construct. This requirement is tested by the principal components analysis of item residuals and by the item fit statistics, which indicate how well each item in a test fits within the underlying construct (Linacre, 2006).

There are many advantages to using the Rasch model over classical test theory (see, for example, Smith, 2003, pp. 1-11, for a summary). First, the Rasch model transforms ordinal-level data to log odd ratios (logits) and places persons and items on an interval scale in a common metric so that the numeric distances between the items on a test and persons in the sample become the actual differences in the level of ability that the test is designed to measure. It also frees the estimates of item and person measures from the distributional properties of test and sample parameters, creating scientifically objective, interval-level measures of ability, which is a prerequisite for conducting parametric statistical analyses.

Second, by estimating the difficulty and discrimination capability of different items and by calibrating the magnitude of differences among test items, Rasch analysis makes it possible to visualize a continuum of construct development and to measure individuals along this continuum. Item-person maps are used to determine how well a test represents the construct and to identify areas for improvement.

Third, the Rasch model makes it possible to detect any aberrant responses in the data (resulting, for example, from carelessness or guessing) and to evaluate the impact of these responses on the person measures. This is done by comparing persons' expected responses with the observed ones and testing the fit of each item and person to the model.

Furthermore, the Rasch model is an effective means for validating the dimensionality of instruments. This is done by performing a principal components analysis of standardized residuals to determine whether the items share any other construct besides the one being measured and by assessing the fit of the item- and person-level data to the measurement model to identify values that depart significantly from those expected by the model. Simulation studies (e.g., Smith & Miao, 1994) have demonstrated that in assessing the dimensionality of a testing instrument designed to produce a unidimensional measure (i.e., in which the majority of items are expected to load on one factor and in which any secondary factor is expected to be highly correlated with the primary factor), Rasch analysis performs better than traditional factor analytic techniques. This finding has led

some researchers (e.g., Smith & Miao) to conclude that when the goal is to create a unidimensional instrument, the Rasch approach is preferable to the traditional factor analytic techniques.

Measurement Models

Two Rasch models were used in the analyses. The models are described below.

Dichotomous Model

In the dichotomous model, the Rasch parameters, item difficulty and person ability, are estimated from the pass-versus-fail proportions for each item and each person (Bond & Fox, 2001). The model is written as follows:

$$\log(P_{ni1}/P_{ni0}) \equiv B_n - D_i$$

where B_n is the ability level (B) of person n, D is an item's (i) difficulty level, P_{ni1} is the probability of person n responding correctly (1) to item i, and P_{ni0} is the probability of the same person responding incorrectly (0) to item i. Thus, the model specifies the relationship between item difficulty and person ability as well as the probability of a correct response to an item by a given person. Item calibration is achieved by identifying each item's location along a latent dimension being measured and the item's location with respect to other items. Both person and item parameters are estimated on the common logit scale. The dichotomous model was used in this study to construct four variables: Phonological Awareness, Short-Term Memory, Decoding Accuracy, and Nonverbal Ability.

Partial Credit Model

The partial credit model was developed by Masters (1982) for use with responses that can be categorized into several ordered categories. This model identifies several intermediate levels of success on items and assigns partial credit for reaching each of these levels. In doing so, it provides a set of individual threshold (*k*) parameters for each item. This model differs from the dichotomous Rasch model in that it contains more than one transition from one category to the next (Bode, 2004). The model is written as follows (Bond & Fox, 2001):

$$\log(P_{nik}/1 - P_{nik}) = B_n - D_{ik}$$

where B_n is the ability level (*B*) of person *n*, *D* is an item's (*i*) difficulty level at threshold *k*, and P_{nik} is the probability of person *n* responding correctly to item *i* at threshold *k*. In this model, responses to an item are classified into ordered categories, with higher-numbered categories representing more of the underlying latent trait measured by the item (Morrison, 1996). This model was used in this study to construct three variables: Listening Comprehension, Reading Comprehension, and Vocabulary.

Analysis Procedures

Rasch analyses were performed using WINSTEPS Version 3.64.2 (Linacre, 2006). The cognitive, linguistic, and decoding measures for which dichotomous or partial credit data were obtained were assumed to reflect seven latent variables: Phonological Awareness (PHAW), Short-Term Memory (STM), Decoding

Accuracy (DACC), Listening Comprehension (LISCOM), Reading Comprehension (READCOM), Vocabulary (VOC), and Nonverbal Ability (IQ). PHAW was represented by the five phonological tests, rhyme detection, initial oddity, final oddity, phoneme segmentation, and phoneme deletion; STM was represented by the two short-term memory tests, nonword repetition and forward digit span; DACC was represented by the accuracy measures on the two reading tests, word reading and nonword reading; LISCOM was represented by the listening comprehension test; READCOM was represented by the reading comprehension test; VOC was represented by the Vocabulary Subtest of the *WISC* (Philimonenko & Timofeev, 2001); and IQ was represented by the Matrices Subtest of the *KBIT-2* (Kaufman & Kaufman, 2004). The obtained scores were rescaled into CHIPS (Linacre, 2006) for SPSS analyses so that all values were positive. In this transformation, 1 logit = 4.55 CHIPS.

Item difficulty and person ability estimates were derived and the hypothetical structure of the constructed variables was examined using Rasch output tables, diagnostic tools, and item-person maps. A three-step procedure was employed to assess dimensionality, as suggested by Linacre (2006). First, the output was examined for negative point-measure correlations for obvious off-dimension behavior. Second, idiosyncratic response patterns were diagnosed using fit statistics and misfitting items and persons were inspected to determine possible causes of the misfit and decide on the best course of action. A principal

components analysis without rotation of standardized residuals was then performed to detect departures from the unidimensionality requirement.

Finally, the items in the PHAW test were examined to identify a hierarchy of item difficulty. The items were hypothesized to form a continuum reflecting Adams' (1990) hierarchy of difficulty, from the cognitively easier rhyme detection items to the more cognitively demanding initial and final oddity, phoneme segmentation, and, finally, phoneme deletion items. Additionally, three predictions were made regarding item difficulty: (a) Items with a simpler syllable structure were expected to be easier than items with a more complex syllable structure; (b) items involving the manipulation of single sounds were expected to be easier than items involving the manipulation of sound clusters (i.e., consonant blends); and (c) items requiring the processing of rhymes were expected to be easier than those requiring the processing of phonemes.

Reliability

The Rasch model provides several estimates of reliability for both persons and items. The person reliability estimate, R_p, (or person separation reliability) is analogous in interpretation to Cronbach alpha, the conventional index of internal consistency (Sheridan & Puhl, 1996, p. 26). In the Rasch model, R_p is the percentage of observed variance that is reproducible (i.e., that is not due to measurement error), and it indicates how well the persons in the sample can be differentiated on the measured variable (Lunz et al., 1996, p. 104). Another

estimate of test reliability is the person separation index, G_p, which shows the spread of persons on the measured variable in terms of units of the measurement error of the scale scores. Mathematically, G_p is estimated as the adjusted person standard deviation divided by the average measurement error, which is defined as "that part of the total variance that is not accounted for by the Rasch model" (Bond & Fox, p. 207). Larger values represent higher reliability.

Similarly, item separation reliability R_i and item separation index G_i are estimates of the reliability of the ordering of items along the measured variable and the replicability of item placement within the hierarchy of items across other samples (Fox & Jones, 1998, p. 36). These indices indicate how adequately the items define the construct under investigation. Acceptable values for Rasch reliability indices are similar to those for Cronbach alpha, with values above .90 being excellent, values between .80 and .90 being very good, and values between .70 and .80 being moderately good.

In this study, item and person separation indices were further transformed into strata statistics, H_p and H_i, using the following formula (Bond & Fox, 2001, p. 207): $H = (4G + 1)/3$. This was done to separate the persons and items into statistically distinct person ability and item difficulty strata using three standard errors to define each stratum.

The Rasch reliability, separation, and strata indices are shown in Table 3. All the item separation reliabilities are in the .81-.98 range (and many are in the high .90s range), indicating that the items in each instrument create a well-defined

variable (Smith, 2003). Most of the person separation reliabilities are somewhat lower, but are still in the acceptable range, with most values being good or very good. These values indicate that the created scales discriminate between the persons well. The only exception is LISCOM (R_p= .64), which had the smallest number of items (n = 15). The ordering of persons on this test is clearly not as reliable as the ordering of persons on the other tests. The Rasch separation indices follow a similar pattern, with person indices being lower than item indices. Here, person separation indices range from a low of 1.34 for the LISCOM variable to a high of 3.17 for the IQ variable, and item separation indices range from 2.10 for the DACC variable to 6.44 for the IQ variable.

Table 3. *Reliability, Separation, and Strata Indices for the Constructed Variables*

Variable	R_p	G_p	H_p	R_i	G_i	H_i
Phonological awareness	.88	2.73	3.97	.93	3.65	5.20
Short-term memory	.80	2.00	3.00	.97	5.80	8.06
Decoding accuracy	.85	2.39	3.52	.81	2.10	3.13
Listening comprehension	.64	1.34	1.78	.97	5.29	7.05
Reading comprehension	.76	1.76	2.68	.94	4.06	5.74
Vocabulary	.79	1.95	2.93	.95	4.52	6.36
Nonverbal ability (IQ)	.91	3.17	4.56	.98	6.44	8.92

Note. R_p = person reliability; G_p= person separation index; R_i= item reliability; G_i = item separation index; H_p = person strata; H_i= item strata. The values reported are "real" values (i.e., the estimated standard errors of measurement have been adjusted for any misfit in the data, Smith, 2003) obtained with extreme persons and items excluded.

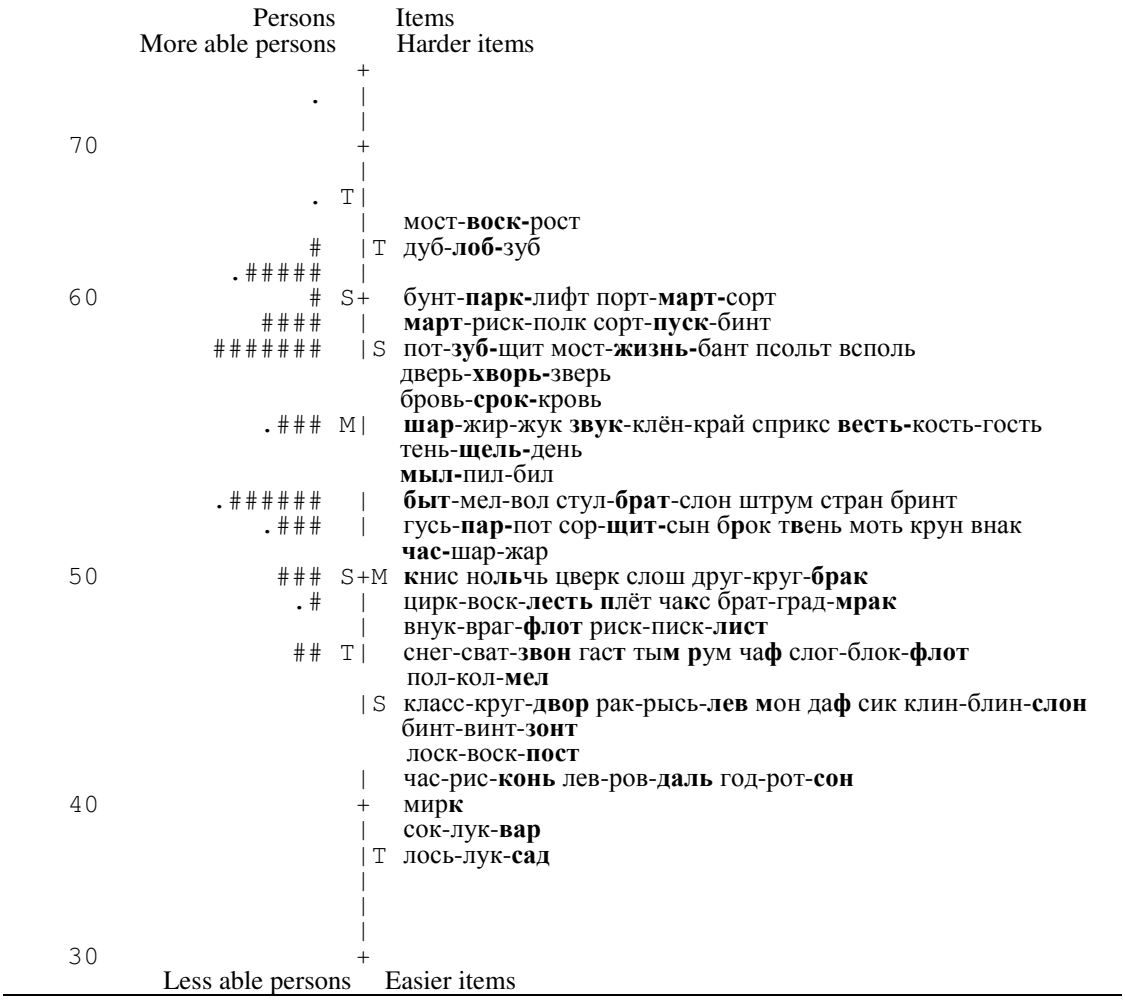
The strata indices in Table 3 indicate the number of person ability strata defined in the sample and the number of item difficulty strata defined in the constructed variables. For person measures, these indices show that depending on the variable, between approximately two and five statistically different strata of ability can be distinguished among the children. For item measures, the strata indices range from approximately three to nine. These high values imply the existence of a large number of statistically distinct areas of difficulty within the hierarchy of items making up each variable.

Item-Person Maps

Figures 2-8 are item-person maps showing the distribution of all the items and persons on the seven constructed variables on the same unidimensional scale. The variables are laid out vertically with the most able persons and most difficult items at the top. The left-hand side locates the person ability measures along each variable. The right-hand side locates the item difficulty measures along each variable. The vertical axis is expressed in CHIPS (1 logit = 4.55 CHIPS). Each # represents two persons; each "." and "x" represent one person. The figures are explained below.

Phonological Awareness

Figure 2 is the item-person map showing the distribution of all the item and person Rasch-scale values on the Phonological Awareness variable. The map



Note. For the rhyme detection, initial oddity, final oddity, and phoneme deletion items, the correct response is highlighted. M = mean, S = 1 *SD*, T = 2 *SD*.

Figure 2. Item-person map of phonological awareness.

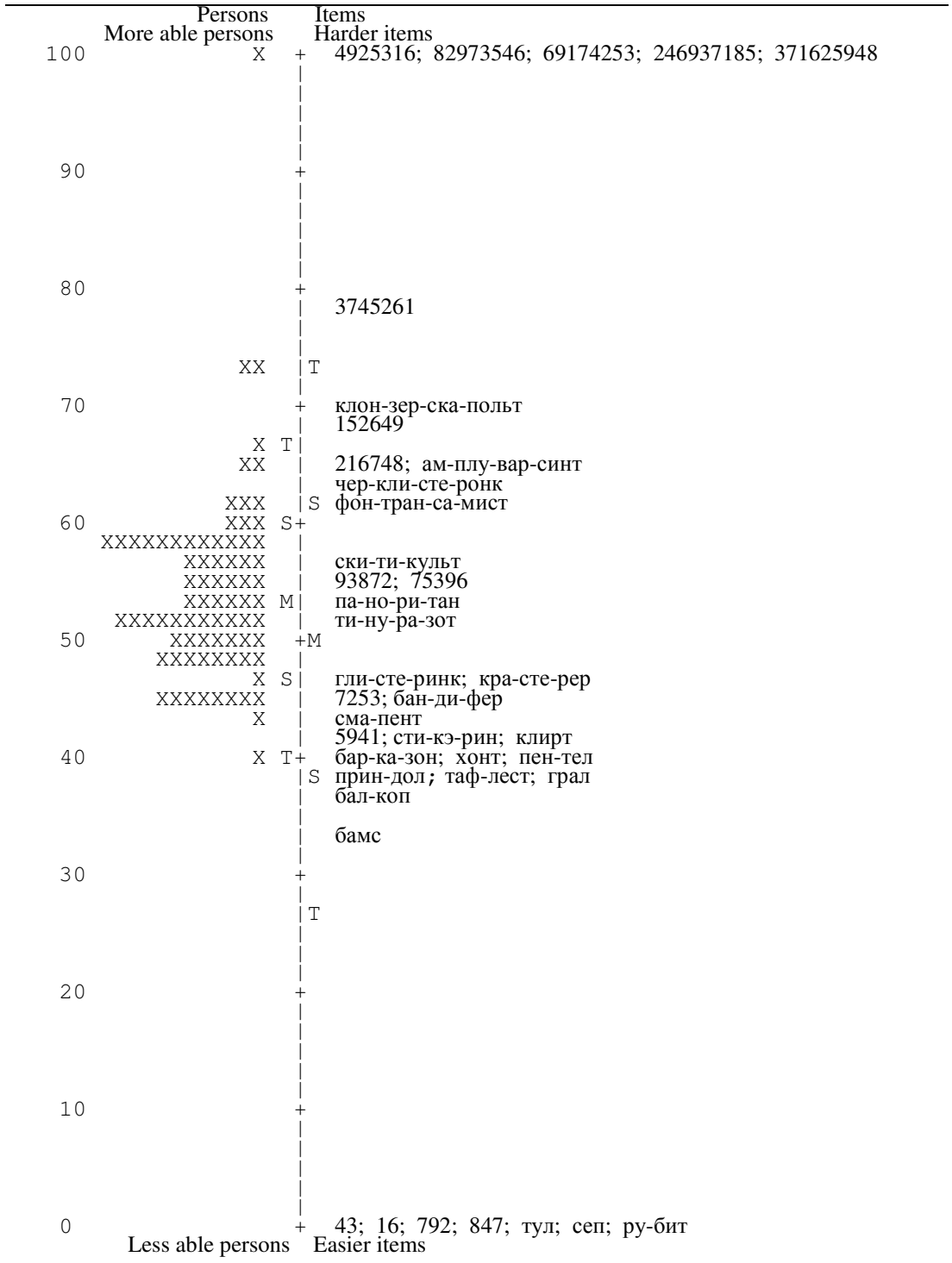
shows the developmental nature of phonological awareness in Russian, with some tasks being more difficult than others. The item calibrations range from 36.4 to 65.6 CHIPS, spanning more than five strata; the person measures range from 44.2 to 72.9 CHIPS, spanning almost four strata. Such a large span indicates that a

variable ranging from more phonological awareness to less phonological awareness was clearly defined.

The test appears to be too easy for this sample, with the mean person ability being almost 1 *SD* above the mean item difficulty and with 21 out of 63 items (or one-third of the test) being more than 2 *SD* lower on the CHIPS scale than the mean ability of the children in the sample. The item-person map shows one small gap in the variable definition, between item RD6 (63.6 CHIPS) and item RD14 (60.5 CHIPS). This gap occurs where 11 of the children are located indicating that differentiation in phonological awareness is somewhat inadequate at this higher level of ability. The figure also shows a considerable redundancy of items, with many items sharing the same difficulty estimate with at least one other item. Most of these items cluster around the mean person ability and 2 *SD* below the mean person ability.

Short-Term Memory

Figure 3 is the item-person map showing the distribution of item and person Rasch-scale values on the Short-Term Memory variable. The item calibrations range from 24.8 to 84.2 CHIPS, spanning more than eight strata; the person measures range from 40.7 to 77.3 CHIPS, spanning three strata. This span indicates that a variable ranging from more short-term memory to less short-term memory was clearly defined.



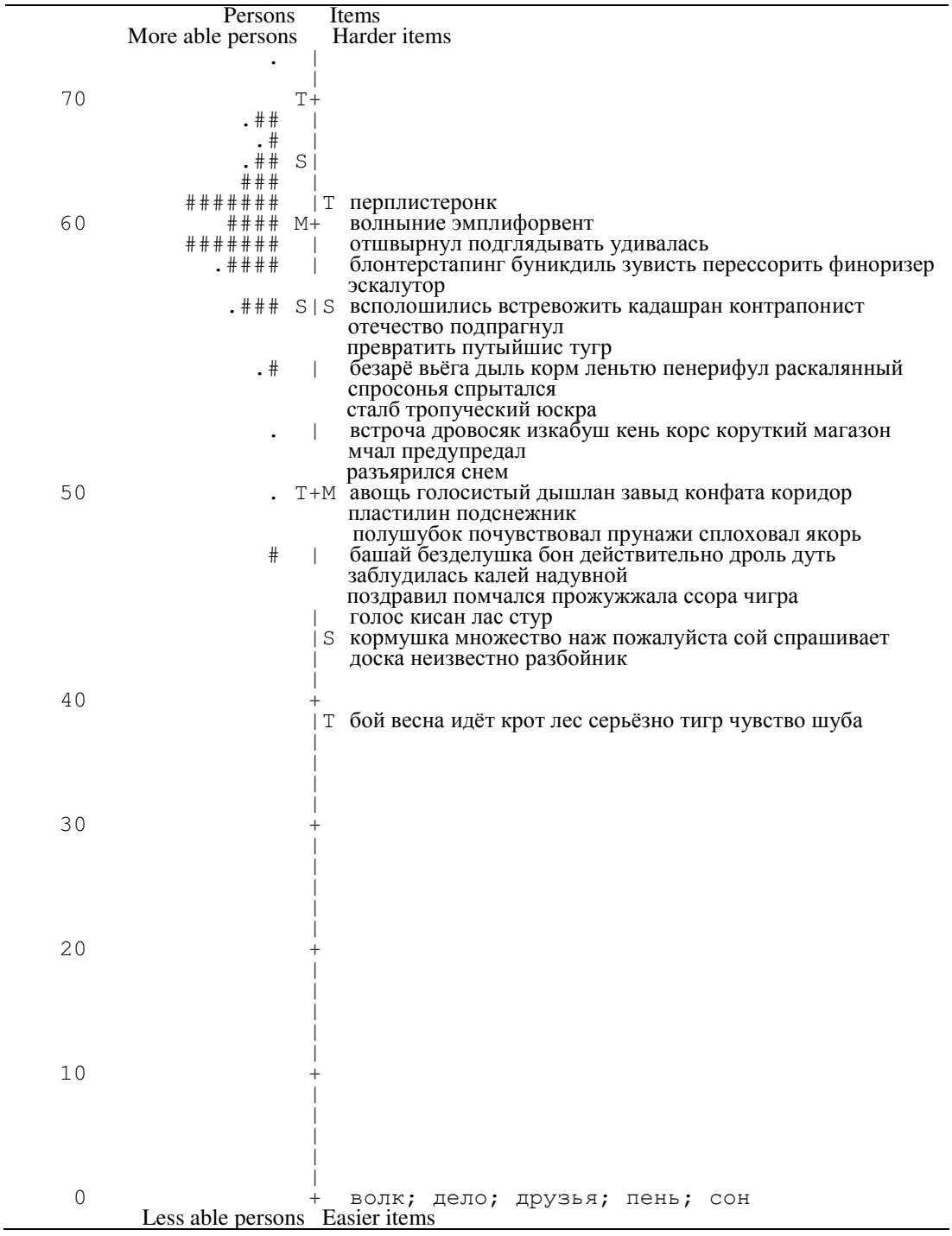
Note. The map shows the actual items used. M = mean, S = 1 *SD*, T = 2 *SD*.

Figure 3. Item-person map of short-term memory.

When the distribution of the sample is compared with the distribution of the items, it is clear that this sample has somewhat more of the STM construct than the mean ability level defined by the items on the test. This is evident in the distance between the two means, which shows that, with the exception of the longest letter and digit sequences, the test is rather easy for this sample. The item-person map also shows two small gaps in places where many persons are located: One between item NR23 (60.9 CHIPS) and item NR16 (57.3 CHIPS) and the other between item NR20 (51.2 CHIPS) and item NR18 (47.2 CHIPS). These gaps occur slightly above and below the mean person ability level, indicating somewhat inadequate differentiation at this level. The figure also shows a considerable redundancy of items, with 14 items being at least 2 *SD* below, and five items being more than 2 *SD* above, the mean person ability.

Decoding Accuracy

Figure 4 is the item-person map showing the distribution of item and person Rasch-scale values on the Decoding Accuracy variable. The item calibrations range from 32.2 to 61.2 CHIPS, spanning more than three strata; the person measures range from 47.6 to 73.2 CHIPS, spanning more than three and a half strata. This span indicates that a variable ranging from more decoding accuracy to less decoding accuracy was clearly defined.



Note. The map shows the actual items used. M = mean, S = 1 *SD*, T = 2 *SD*.

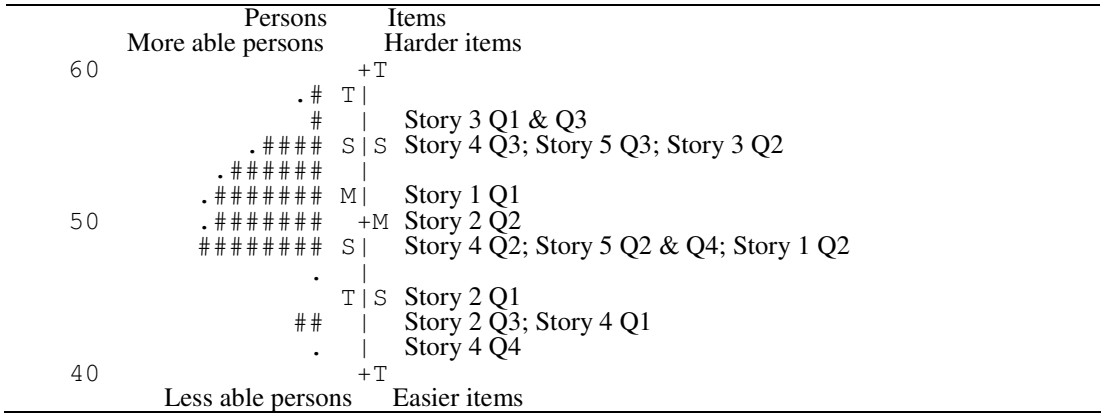
Figure 4. Item-person map of decoding accuracy.

The map shows that the mean for the person measures (left side) is located much higher than the mean for the item calibrations (right side). This test is clearly too easy for this sample: Twenty-seven out of 98 items, or 27.5%, have difficulty estimates below the lowest person ability estimate. In fact, the map shows a floor effect problem on the test: The lower end of the scale has many more items than the middle part, whereas the higher end has no items at all. The mean of item difficulty is located 2 *SD* below the mean of person ability and the most difficult item, a four-syllable nonword with three complex consonant clusters, is located just above the mean person ability level. Overall, the map indicates that the children in the sample found decoding even long and phonologically complex words a relatively easy task. The map also shows a considerable redundancy of items, as many items share a similar difficulty estimate with at least one other item.

Listening Comprehension

Figure 5 shows the distribution of item and person Rasch-scale values on the Listening Comprehension variable, which has three levels of performance: complete success (a score of 2), partial success (a score of 1), and failure (a score of 0). The item calibrations range from 41.7 to 57.4 CHIPS, spanning more than seven strata; the person measures range from 41.4 to 59.1 CHIPS, spanning almost two strata. This span indicates that a variable ranging from more listening comprehension to less listening comprehension was clearly defined.

The item map shows a good match between the test and the sample. The mean of person measures (51.2) is close to the mean of item calibrations (50.0), and most persons are well aligned with the items along the CHIPS scale. There is, however, a small gap in coverage just above the item mean where a relatively large number of children are located. This gap indicates that the test may not be measuring the ability of some of the children as precisely as that of others.



Note. Q = question, M = mean, S = 1 *SD*, T = 2 *SD*. The stories are in Appendix M.

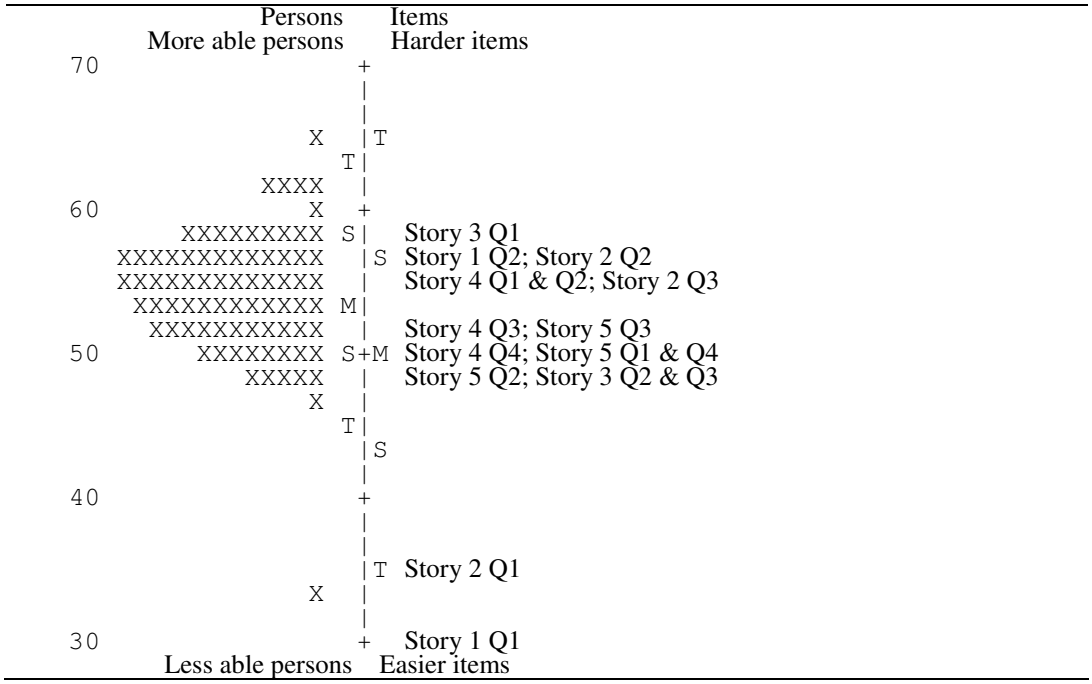
Figure 5. Item-person map of listening comprehension.

Reading Comprehension

Figure 6 is the item-person map showing the distribution of item and person Rasch-scale values on the Reading Comprehension variable, which has three levels of performance: complete success (a score of 2), partial success (a score of 1), and failure (a score of 0). The item calibrations range from 30.8 to 58.2 CHIPS, spanning almost three strata; the person measures range from 32.5 to 65.2 CHIPS,

spanning almost six strata. Such a large span indicates that a variable ranging from more reading comprehension to less reading comprehension was clearly defined.

The item map shows some mismatch between the test and the sample: The mean of item calibrations (50.0) is located approximately 1 *SD* below the mean of person measures (54.1). The test is clearly easy for this sample, although it does capture a significant number of children in the middle-ability range. There are also no items in the high-ability range, which indicates that the test may not be measuring high-ability children precisely.



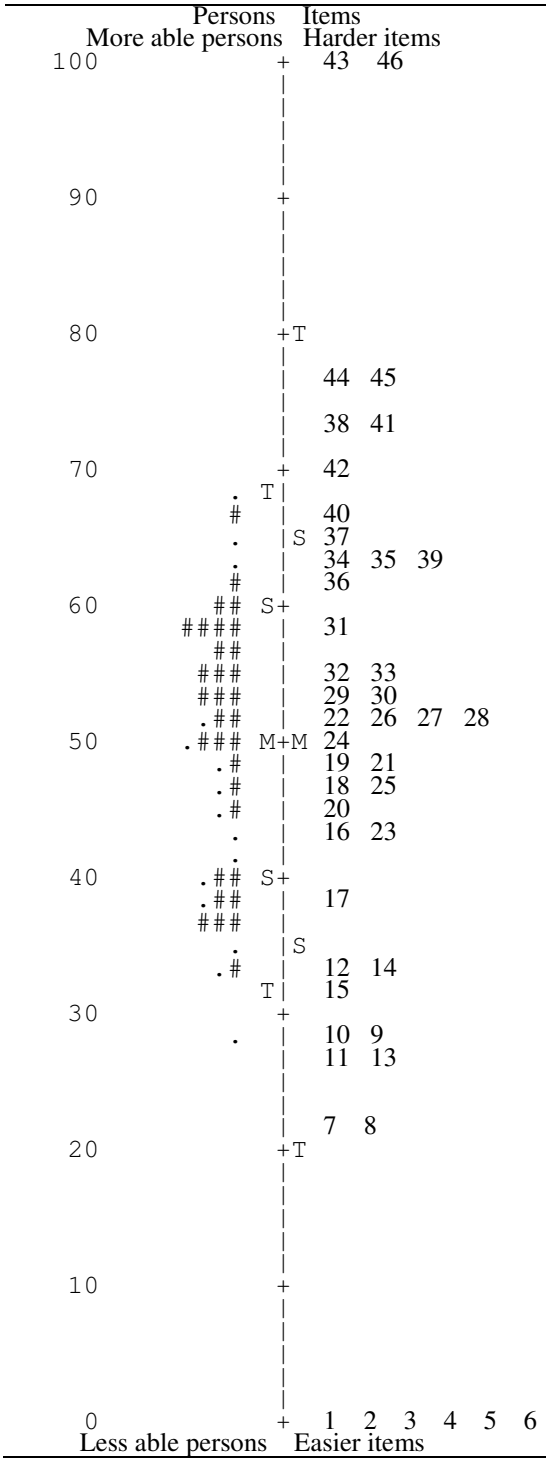
Note. Q = question, M = mean, S = 1 *SD*, T = 2 *SD*. The stories are in Appendix L.

Figure 6. Item-person map of reading comprehension.

Nonverbal Ability

Figure 7 is the item-person map showing the distribution of item and person Rasch-scale values on the IQ variable. The item calibrations range from 16.1 to 82.1 CHIPS, spanning more than four and a half strata; the person measures range from 28.8 to 68.1 CHIPS, spanning nine strata. Such a large span indicates that a variable ranging from more nonverbal ability to less nonverbal ability was clearly defined.

The item map shows a good match between the test and the sample. The mean of person measures is very close to the mean of item calibrations and most persons are located within 1 *SD* of the mean item difficulty. However, gaps in the coverage of the CHIPS scale are apparent in four places, between items 12 and 17 (6.5 CHIPS), between items 16 and 17 (4.8 CHIPS), between items 31 and 32 (3.8 CHIPS), and between items 31 and 36 (3.5 CHIPS). Although these gaps are not very large, they appear in places where many children are located indicating that the precision of person ability estimates is not optimal for some of the children. The figure also shows redundant items at the lowest and highest ends of the scale: Items 1 through 8 are too easy and items 38 and 41 through 46 are too difficult for this sample. The bottom items in particular show a clear floor effect; these items (together with item 9) constitute the easiest section of the Matrices test used to operationalize IQ. In fact, the authors recommend that this section be skipped with older children. The item map confirms that these items are not useful for use with Russian children of age seven and up.



Note. Q = question, M = mean, S = 1 *SD*, T = 2 *SD*.

Figure 7. Item-person map of nonverbal ability.

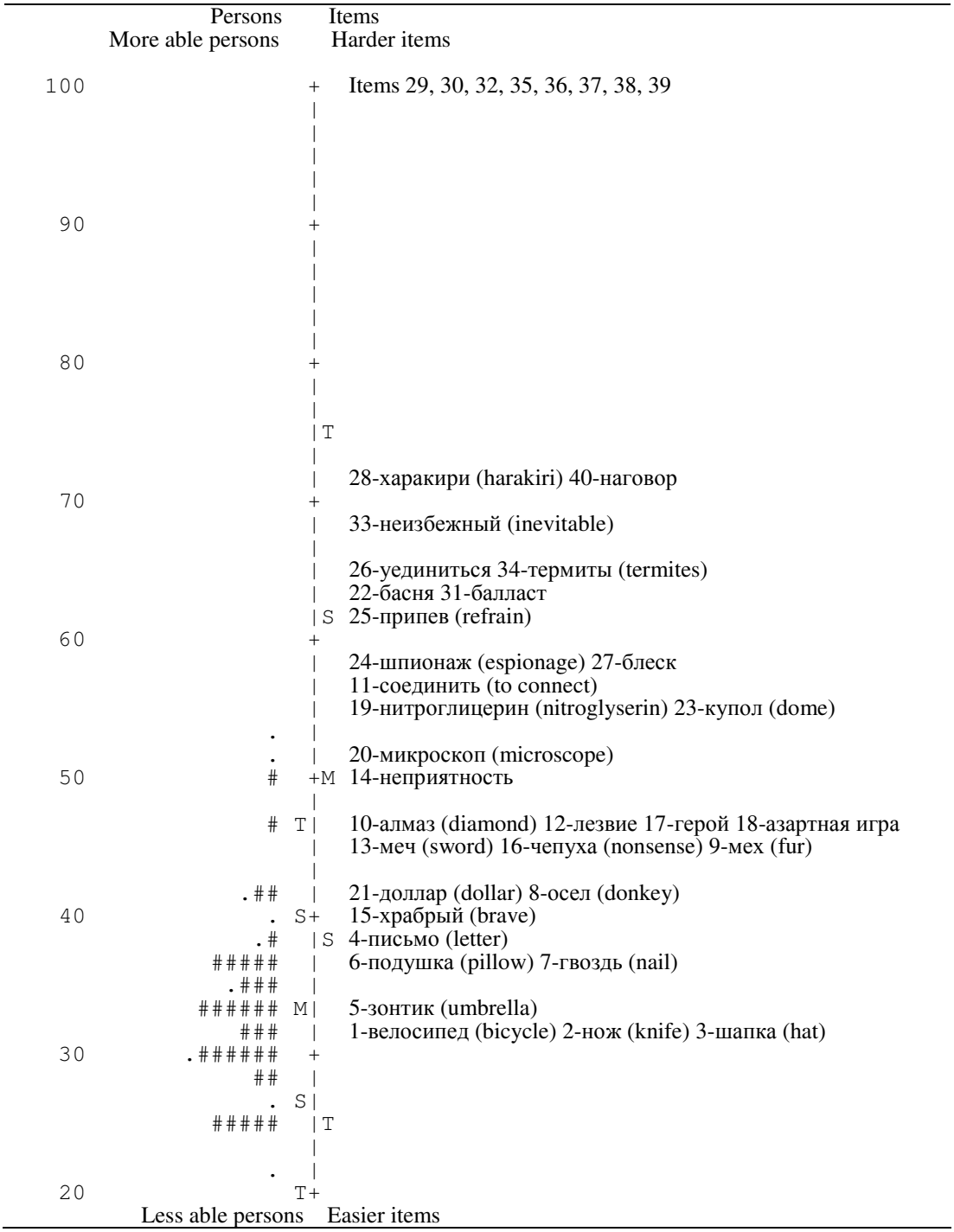
Vocabulary

Figure 8 is the item-person map showing the distribution of item and person Rasch-scale values on the Vocabulary variable, which has three levels of performance: complete success (a score of 2), partial success (a score of 1), and failure (a score of 0). The item calibrations range from 30.9 to 77.5 CHIPS, spanning three strata; the person measures range from 21.8 to 53.1 CHIPS, spanning more than six strata. These values show a discrepancy between the level of difficulty of the items and the level of ability of the children in the sample. In fact, as shown in the figure, the mean item difficulty is more than 2 *SD* above the mean person ability and on average, the children in this sample were able to define only a very small number of items.

The test appears to be particularly inappropriate for children in the lower-to-middle ability range, and although it does a slightly better job for children in the higher ability range, overall, there are very few children in this sample who are measured precisely by this test. The figure also shows a considerable redundancy of items, with more than half not targeting any persons at all.

Point-Measure Correlations

The output was examined for negative point-measure correlations to determine if there were obvious off-dimension responses. According to Linacre (2006), under Rasch conditions, point-measure correlations should be positive so that the item-level scoring accords with the latent variable (p. 298). Four negative correlations



Note. M = mean, S = 1 *SD*, T = 2 *SD*. English translations for some of the items are shown in parentheses.

Figure 8. Item-person map of vocabulary.

were found: for item 1 in DACC and READCOM and for items 7 and 13 in IQ; these items also showed severe misfit to the model. These items were deleted together with other misfitting items (see below) and the remaining items were recalibrated. The recalibration resulted in positive point-measure correlations for all the items.

Fit Statistics

In Rasch analysis, fit statistics are used to determine whether the items in a test measure the same construct and whether the persons in a sample respond in a manner consistent with that predicted by the model. Item fit indicates the extent to which an item discriminates between persons with high ability and persons with low ability in a way that is consistent with other items in the test. Person fit indicates the extent to which the person's pattern of responses is consistent with the way the items are used by the other persons (Bond & Fox, 2001).

The Rasch model provides two indicators of fit: the infit statistic and the outfit statistic. The infit statistic is an estimate of the degree of fit between the actual observations and the Rasch-modeled expectations weighted by the information in the data; this statistic places more emphasis on unexpected responses near a person's or item's measure. The outfit statistic is an unweighted estimate of the fit and it places more emphasis on unexpected responses far from a person's or item's measure. Both types of statistic are available in two forms, unstandardized as mean squares and standardized as *t*.

An important question in Rasch analysis is how to determine which persons and which items misfit the Rasch model. Although simulation studies have provided some guidelines, the Rasch literature is replete with conflicting recommendations as to which indices to use when making decisions about fit. For example, Fox and Jones (1998, p. 41) and Bond and Fox (2001) argue that the current practice is to use mean square, rather than standardized values and to focus on infit statistics because the outfit statistic is more sensitive to outliers and a few random responses may produce high outfit mean square values. In a similar vein Linacre (2006) has argued that poor infit mean square values pose a greater threat to validity than poor outfit mean squares because they are influenced by response patterns that may reflect systematic bias. In contrast, outfit mean squares pose less threat to validity because they reflect extreme, or aberrant, responses. On the other hand, Smith (2000) has argued that standardized unweighted fit statistics are a better choice for assessing fit to the measurement model because they have more stable Type 1 error rates and more consistent distributional properties vis-à-vis sample size than the mean square statistics. Because of the lack of consensus concerning fit criteria, all four types of fit statistic were examined in this study.

For the mean square statistic, the critical values chosen for flagging misfitting items were .5 and 1.7, as recommended by Linacre and Wright (1994); the critical values chosen for flagging misfitting persons were .3 and 2.0. The slightly greater range for acceptable person fit was used because the sample consisted of young children unfamiliar with cognitive testing. For the standardized

fit statistic, values greater than ± 2.0 for both items and persons were interpreted as having less compatibility with the model than expected, with negative values interpreted as indicating overfit, or the closeness of responses to the deterministic Guttman pattern. These values correspond approximately to a Type 1 error rate of 5% (Smith, 1991, 2004, p. 580). After misfitting items and persons were identified, the most unexpected responses for these items and persons were carefully inspected to determine the nature of the misfit. The quality of the misfitting items and the consistency of the misfitting persons were then considered before making a final decision.

Item Fit

An examination of item fit statistics revealed a number of misfitting items on most of the variables: There were eight misfitting items (5%) in PHAW (rhyme detection items 2, 8, 11, and 17; initial oddity items 1 and 10; final oddity items 7 and 9), four misfitting items (10%) in STM (nonword repetition items 3, 5, 6, and 9), five misfitting items (3%) in DACC (word reading items 1, 29, 44, and 47), two misfitting items (12%) in READCOM (Story 1 question 1 and Story 2 question 3), 19 misfitting items (41%) in IQ (items 7, 8, 11-15, 17, 18 20, 23, 32, 36, 39-42, 44, and 45), and 19 misfitting items (45%) in VOC (items 1-3, 5, 6, 8, 10, 18, 19, 21, 22, 26-28, 31, 33, 34, 40). Eleven of the misfitting items on the IQ test and 13 of the misfitting items on the VOC test overfit the model, with values only slightly below the cut-off point. All of these items misfit because of just one unexpected

response; thus, these items were retained. All the other misfitting items were deleted and the data were recalibrated, as suggested by Smith (2003).

Tables 4-10 present a summary of the individual statistics for each of the remaining items for each of the variables. The tables list the entry order of the items, the raw score, the CHIPS measure, the error of estimate for the measure, and the fit statistics. Items with a perfect score are not shown.

After the deletion of the misfitting items, an improved calibration was obtained for all the variables, with only two items exceeding the cut-off values in PHAW, one item exceeding the cut-off values in STM, three items exceeding the cut-off values in DACC, two items exceeding the cut-off values in RC, four items exceeding the cut-off values in IQ, and one item exceeding the cut-off value in VOC. Two items on the IQ test (items 9 and 10) had mean square values of 3.90 and 9.90, respectively. All the other items with fit statistics exceeding the critical values had mean squares in the 1.8-2.2 range and standardized values in the 2.1-2.8 range, and most misfit on just one of the fit statistics.

To diagnose the nature of the misfit, the most unexpected responses for the misfitting items were carefully examined. The examination revealed that in most cases, the misfit was caused by unexpected responses from just a few children (about 5%) who tended to miss easy items. For example, the misfit of IQ items 9 and 10, the most severely misfitting items, was caused by random responses from

Table 4. *Rasch Analysis of Phonological Awareness: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
RD1	42.40	73.00	2.00	0.83	-0.46	0.57	-0.53
RD3	52.80	51.00	1.20	1.18	1.52	1.23	1.09
RD4	56.10	40.00	1.20	1.13	1.27	1.24	1.41
RD5	44.70	70.00	1.70	1.00	0.09	1.24	0.59
RD6	64.60	15.00	1.40	1.05	0.35	1.04	0.24
RD7	45.40	69.00	1.70	1.35	1.41	1.93	1.59
RD9	48.40	63.00	1.40	1.23	1.32	1.01	0.14
RD10	50.40	58.00	1.30	1.27	1.81	1.28	1.01
RD12	43.30	72.00	1.90	0.91	-0.22	0.56	-0.64
RD13	43.30	72.00	1.90	1.09	0.40	1.12	0.40
RD14	61.40	23.00	1.30	1.12	0.89	1.60	2.11
RD15	55.50	42.00	1.20	0.85	-1.55	0.74	-1.62
RD16	43.30	72.00	1.90	1.08	0.37	0.97	0.16
RD18	48.00	64.00	1.40	1.22	1.19	1.41	1.07
IO2	52.10	53.00	1.20	0.94	-0.49	0.86	-0.57
IO3	43.30	72.00	1.90	0.93	-0.13	0.64	-0.44
IO4	52.80	51.00	1.20	1.07	0.64	1.04	0.27
IO5	55.80	41.00	1.20	0.91	-0.93	0.90	-0.57
IO6	45.40	69.00	1.70	1.33	1.33	1.73	1.33
IO7	46.50	67.00	1.60	1.01	0.10	0.88	-0.11
IO8	53.70	48.00	1.20	0.96	-0.31	0.89	-0.53
IO10	44.00	71.00	1.80	0.97	0.00	0.53	-0.79
FO1	38.80	76.00	2.80	1.05	0.26	1.49	0.77
FO2	41.40	74.00	2.20	1.02	0.18	0.73	-0.16
FO3	54.30	46.00	1.20	1.12	1.10	1.16	0.88
FO4	57.90	34.00	1.20	1.18	1.69	1.28	1.54
FO5	41.40	74.00	2.20	1.04	0.24	0.95	0.18
FO6	57.30	36.00	1.20	0.97	-0.22	0.93	-0.35
FO8	59.80	28.00	1.20	0.97	-0.21	0.89	-0.49
FO10	59.50	29.00	1.20	1.14	1.26	1.31	1.48
PS1	44.00	71.00	1.80	0.85	-0.47	0.83	-0.12
PS2	51.80	54.00	1.20	1.20	1.54	1.37	1.49
PS3	45.40	69.00	1.70	0.74	-1.13	0.63	-0.66
PS4	51.80	54.00	1.20	0.78	-1.91	0.66	-1.58
PS5	50.00	59.00	1.30	0.79	-1.51	0.65	-1.27
PS6	50.40	58.00	1.30	0.85	-1.05	0.79	-0.72

(Table 4 continues)

(Table 4 continued)

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
PS7	53.40	49.00	1.20	0.86	-1.30	0.74	-1.38
PS8	53.10	50.00	1.20	0.80	-1.86	0.70	-1.56
PS9	57.90	34.00	1.20	0.85	-1.49	0.77	-1.36
PS10	53.40	49.00	1.20	0.89	-0.99	1.05	0.32
PS11	53.70	48.00	1.20	0.98	-0.13	0.99	0.00
PS12	58.20	33.00	1.20	1.19	1.73	1.16	0.89
PS13	55.80	41.00	1.20	1.02	0.25	1.15	0.90
PD1	44.00	71.00	1.80	0.76	-0.87	0.44	-1.05
PD2	45.40	69.00	1.70	0.80	-0.83	0.65	-0.61
PD3	51.10	56.00	1.30	0.84	-1.27	0.70	-1.23
PD4	48.80	62.00	1.40	1.01	0.11	1.32	0.96
PD5	51.50	55.00	1.30	1.00	0.07	1.06	0.33
PD6	51.80	54.00	1.20	0.88	-0.95	0.79	-0.89
PD7	49.60	60.00	1.30	0.88	-0.81	0.74	-0.82
PD8	50.80	57.00	1.30	1.03	0.23	0.93	-0.18
PD9	44.00	71.00	1.80	0.99	0.05	0.74	-0.31
PD10	44.70	70.00	1.70	0.89	-0.35	0.97	0.12
PD11	45.40	69.00	1.70	0.89	-0.41	0.87	-0.10
PD12	40.20	75.00	2.40	1.00	0.15	0.65	-0.20

Note. RD = rhyme detection; IO= initial oddity; FO = final oddity; PS = phoneme segmentation; PD = phoneme deletion. See Appendices A-E for the actual items used. Mnsq = mean squares; zstd = standardized values.

just two low-ability children. Temporarily removing these children from item analysis resulted in an excellent fit for these items. The recalibration also produced a small number of overfitting items, i.e., items with fit values below the cut-off point, but an inspection of response patterns for these items revealed that all of them were caused by just one or two overly predictable responses. All of the recalibrated items were therefore retained for analysis.

Table 5. *Rasch Analysis of Short-Term Memory: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
NW4	35.80	73.00	2.20	1.01	0.16	0.56	-0.34
NW8	37.60	71.00	1.90	1.11	0.47	0.84	0.01
NW10	35.80	73.00	2.20	0.91	-0.13	0.41	-0.64
NW11	35.80	73.00	2.20	1.18	0.58	1.91	1.14
NW12	40.30	67.00	1.60	0.86	-0.61	0.57	-0.71
NW13	38.40	70.00	1.80	0.83	-0.59	0.50	-0.69
NW14	42.40	63.00	1.50	0.92	-0.44	0.95	0.06
NW15	39.70	68.00	1.70	0.99	0.05	0.67	-0.43
NW16	55.60	26.00	1.30	1.23	1.56	1.25	0.94
NW17	44.10	59.00	1.40	0.88	-0.84	0.63	-0.80
NW18	44.90	57.00	1.30	1.06	0.51	1.00	0.14
NW19	50.80	40.00	1.20	0.92	-0.63	0.87	-0.47
NW20	49.10	45.00	1.20	0.91	-0.71	0.79	-0.70
NW21	63.40	10.00	1.80	1.05	0.27	0.98	0.18
NW22	62.10	12.00	1.70	0.82	-0.78	0.44	-1.12
NW23	59.40	17.00	1.50	1.07	0.44	0.83	-0.31
NW24	68.00	5.00	2.40	0.74	-0.62	0.37	-0.54
DS5	39.70	68.00	1.70	1.10	0.50	1.03	0.25
DS6	41.90	64.00	1.50	1.20	1.14	1.03	0.23
DS7	53.80	31.00	1.30	0.97	-0.18	1.37	1.47
DS8	53.50	32.00	1.30	1.26	1.83	1.36	1.45
DS9	66.80	6.00	2.20	0.84	-0.36	0.43	-0.52
DS10	64.10	9.00	1.90	1.00	0.08	0.66	-0.33
DS11	77.00	1.00	4.80	1.17	0.48	0.52	-0.06

Note. NW = nonword repetition; DS = digit span. See Appendices F and G for the actual items used. Mnsq = mean squares; zstd = standardized values.

Person Fit

The recalibration of the items also improved person fit for most of the variables, although a certain number of misfitting persons still remained. For all the

Table 6. *Rasch Analysis of Decoding Accuracy: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
1	45.70	73.00	2.20	0.87	-0.22	0.70	-0.27
2	47.50	71.00	1.90	0.96	-0.03	0.85	-0.10
3	51.20	65.00	1.50	1.00	0.05	1.06	0.27
4	51.20	65.00	1.50	1.11	0.61	1.17	0.57
5	46.70	72.00	2.00	0.73	-0.76	0.34	-1.30
6	49.00	69.00	1.70	0.87	-0.45	0.81	-0.29
7	48.30	70.00	1.80	0.85	-0.49	1.19	0.52
8	48.30	70.00	1.80	0.88	-0.35	0.71	-0.49
9	46.70	72.00	2.00	0.78	-0.61	0.43	-1.02
10	48.30	70.00	1.80	0.85	-0.48	0.68	-0.55
11	52.60	62.00	1.40	0.90	-0.52	0.66	-1.24
12	50.20	67.00	1.60	1.05	0.30	1.06	0.29
13	53.80	59.00	1.30	1.06	0.43	1.12	0.55
14	52.20	63.00	1.40	0.92	-0.39	0.72	-0.91
15	50.70	66.00	1.50	0.94	-0.23	0.98	0.06
16	54.90	56.00	1.30	1.02	0.21	0.95	-0.17
17	54.50	57.00	1.30	1.00	0.04	1.11	0.56
18	56.80	50.00	1.20	0.88	-1.13	0.79	-1.31
19	53.00	61.00	1.40	1.00	0.04	1.02	0.15
20	57.10	49.00	1.20	1.00	-0.01	1.02	0.15
21	54.90	56.00	1.30	0.87	-0.99	0.74	-1.28
22	61.50	34.00	1.10	0.84	-1.72	0.78	-1.61
23	59.50	41.00	1.10	1.03	0.35	1.02	0.21
24	57.10	49.00	1.20	0.96	-0.40	0.92	-0.49
25	46.70	72.00	2.00	0.80	-0.51	0.68	-0.40
26	45.70	73.00	2.20	0.87	-0.24	0.70	-0.27
27	51.20	65.00	1.50	0.78	-1.13	0.58	-1.33
28	54.10	58.00	1.30	0.92	-0.56	0.88	-0.47
29	54.50	57.00	1.30	0.94	-0.42	1.01	0.13
30	53.00	61.00	1.40	0.83	-1.04	0.99	0.05
31	50.20	67.00	1.60	0.87	-0.52	0.89	-0.15
32	53.40	60.00	1.30	1.10	0.69	1.22	0.89
33	49.60	68.00	1.70	0.88	-0.44	0.66	-0.77
34	53.80	59.00	1.30	0.94	-0.39	0.82	-0.68
35	56.80	50.00	1.20	0.99	-0.04	1.05	0.36
36	51.20	65.00	1.50	0.87	-0.60	0.73	-0.74

(Table 6 continues)

(Table 6 continued)

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
37	51.20	65.00	1.50	0.97	-0.11	1.15	0.53
38	50.20	67.00	1.60	0.89	-0.45	0.85	-0.26
39	51.70	64.00	1.50	1.01	0.10	1.10	0.41
40	51.20	65.00	1.50	0.83	-0.84	0.65	-1.05
41	53.40	60.00	1.30	1.01	0.11	0.93	-0.18
42	54.90	56.00	1.30	1.06	0.51	1.23	1.07
43	59.50	41.00	1.10	0.94	-0.62	0.93	-0.52
44	58.00	46.00	1.20	0.96	-0.37	0.93	-0.46
45	56.20	52.00	1.20	0.89	-1.01	0.79	-1.23
46	52.60	62.00	1.40	0.90	-0.52	1.09	0.38
47	51.20	65.00	1.50	0.97	-0.06	0.90	-0.19
48	53.40	60.00	1.30	1.07	0.50	1.16	0.65
51	37.70	77.00	4.70	1.04	0.36	0.60	0.08
53	37.70	77.00	4.70	1.06	0.38	1.16	0.55
54	37.70	77.00	4.70	1.02	0.33	0.37	-0.21
55	43.10	75.00	2.80	0.86	-0.11	0.33	-0.72
56	37.70	77.00	4.70	1.00	0.32	0.31	-0.30
57	49.60	68.00	1.70	1.06	0.32	0.80	-0.37
59	37.70	77.00	4.70	1.00	0.32	0.31	-0.30
60	48.30	70.00	1.80	1.08	0.36	1.05	0.26
61	50.20	67.00	1.60	1.24	1.06	1.24	0.68
62	45.70	73.00	2.20	1.14	0.49	1.20	0.50
63	50.20	67.00	1.60	1.18	0.84	1.01	0.15
64	43.10	75.00	2.80	1.09	0.34	0.99	0.30
65	45.70	73.00	2.20	1.09	0.35	0.71	-0.25
66	49.00	69.00	1.70	1.27	1.05	1.40	0.93
67	48.30	70.00	1.80	1.26	0.95	1.68	1.30
68	58.00	46.00	1.20	1.18	1.74	1.37	2.31
69	37.70	77.00	4.70	1.05	0.37	0.79	0.26
71	48.30	70.00	1.80	1.22	0.81	1.42	0.91
72	51.20	65.00	1.50	1.11	0.61	1.21	0.68
73	52.60	62.00	1.40	1.11	0.69	1.01	0.15
74	37.70	77.00	4.70	1.02	0.33	0.37	-0.21
76	47.50	71.00	1.90	1.07	0.32	0.91	0.01
78	37.70	77.00	4.70	1.04	0.36	0.60	0.08
79	43.10	75.00	2.80	1.00	0.17	0.90	0.19
80	54.10	58.00	1.30	1.01	0.14	0.97	-0.05

(Table 6 continues)

(Table 6 continued)

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
81	49.00	69.00	1.70	1.10	0.48	1.64	1.33
82	48.30	70.00	1.80	1.18	0.69	0.96	0.09
83	45.70	73.00	2.20	1.00	0.13	0.59	-0.49
84	49.60	68.00	1.70	1.14	0.65	1.90	1.82
85	47.50	71.00	1.90	0.98	0.05	2.29	1.93
86	56.50	51.00	1.20	0.98	-0.12	0.98	-0.04
87	50.20	67.00	1.60	1.26	1.14	1.63	1.46
88	50.70	66.00	1.50	1.19	0.90	1.08	0.33
89	50.70	66.00	1.50	1.19	0.90	0.99	0.11
90	50.70	66.00	1.50	0.86	-0.64	0.63	-1.02
91	54.90	56.00	1.30	0.96	-0.26	0.86	-0.60
93	55.20	55.00	1.20	1.21	1.57	1.29	1.38
94	45.70	73.00	2.20	1.18	0.57	1.23	0.54
96	52.20	63.00	1.40	1.29	1.52	1.67	1.90
97	48.30	70.00	1.80	1.23	0.86	1.26	0.65

Note. Items 1-48 are nonwords; items 51-97 are words; see Appendices J and K for the actual items used. Mnsq = mean squares; zstd = standardized values.

variables except IQ, there were between four and six persons (5-7%) with fit values exceeding the cut-off values; for IQ, there were 10 such persons, or 12%. This was an improvement, however, compared to the 17% before the item recalibration. An inspection of the most unexpected response patterns showed that most of the misfitting children responded in an inconsistent manner to a small number of items (1-10%, depending on the test).

On all of the tests except DACC there were also a small number of children with fit statistics below the cut-off point, indicating unexpected consistency in responses. An inspection of the response patterns for these children revealed that most of the overfit was caused by unexpected responses to a very small number of

items, typically just one or two. Because the overfit was caused by just a few arbitrary responses and because the amount of overfit was relatively small, these children were not considered a threat to the validity of the measurement and were retained for analysis.

Table 7. *Rasch Analysis of Listening Comprehension: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
S1-1	51.10	82.00	0.60	1.20	1.68	1.20	0.96
S1-2	49.00	92.00	0.90	0.91	-0.62	0.90	-0.67
S2-1	45.70	134.00	0.80	0.88	-0.47	0.58	-0.81
S2-2	49.60	98.00	0.70	0.92	-0.61	0.90	-0.42
S2-3	43.80	143.00	1.00	1.31	1.02	1.45	0.85
S3-1	57.40	37.00	0.90	0.89	-0.69	0.79	-1.17
S3-2	55.80	34.00	0.80	0.92	-0.46	1.03	0.19
S3-3	57.30	34.00	0.90	0.90	-0.55	0.87	-0.55
S4-1	44.10	126.00	1.00	1.18	1.06	1.12	0.64
S4-2	48.20	108.00	0.70	0.93	-0.49	0.88	-0.59
S4-3	54.30	55.00	0.80	1.25	1.79	1.18	1.19
S4-4	41.70	149.00	1.30	0.81	-0.39	0.48	-0.64
S5-1	48.40	112.00	0.70	1.26	1.70	1.54	1.67
S5-2	54.40	61.00	0.90	0.96	-0.22	0.97	-0.16
S5-3	49.10	100.00	0.70	0.82	-1.49	0.76	-1.50

Note. S1 = Story 1, S2 = Story 2, S3 = Story 3, S4 = Story 4, S5 = Story 5. The second digit under "Item" denotes the question number. See Appendix M for the actual questions used. Mnsq = mean square values; zstd = standardized values.

Overall, after the recalibration of the items, the fit statistics for all the items and most of the children were acceptable, indicating that the data fit the model well. There were several items and persons with misfitting values, most notably on the

IQ test. However, the misfit in most cases involved a small number of items and persons and was caused by just a few random responses. It was therefore unlikely that either the items or the persons in the study exhibited any systematic bias.

Table 8. *Rasch Analysis of Reading Comprehension: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
S1-2	56.90	52.00	0.80	0.96	-0.25	1.00	0.08
S2-1	31.90	155.00	4.60	1.00	0.32	0.50	-0.08
S2-2	56.50	51.00	0.80	1.18	1.15	1.84	2.81
S3-1	57.60	103.00	1.20	1.21	1.70	1.24	1.34
S3-2	47.60	126.00	0.90	0.97	-0.12	0.85	-0.21
S3-3	47.00	127.00	0.90	1.10	0.55	1.59	1.44
S4-1	54.50	70.00	0.80	1.20	1.47	1.28	1.69
S4-2	54.10	72.00	0.70	0.94	-0.46	0.96	-0.18
S4-3	50.60	104.00	0.70	0.91	-0.60	0.71	-1.03
S4-4	48.80	117.00	0.80	0.82	-1.15	0.64	-1.14
S5-1	48.20	121.00	0.80	0.83	-1.00	0.58	-1.16
S5-2	47.90	126.00	0.80	1.08	0.48	1.91	1.47
S5-3	50.10	112.00	0.70	0.91	-0.55	0.83	-0.09
S5-4	48.10	125.00	0.80	0.68	-1.89	0.46	-0.93

Note. S1 = Story 1, S2 = Story 2, S3 = Story 3, S4 = Story 4, S5 = Story 5. The second digit under "Item" denotes the question number. See Appendix M for the actual questions used. Mnsq = mean square values; zstd = standardized values.

A somewhat better calibration could be obtained by dropping the misfitting persons and recalibrating the data; however, as Smith (2003, p. 29) has pointed out, in this case, the changes in the item calibrations would be much smaller than the standard errors of estimate for the items and there would be no change in the definition of the variable. This would also dramatically reduce the number of

Table 9. *Rasch Analysis of Nonverbal Ability: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
8	6.20	78.00	5.10	0.53	-0.42	0.03	-3.23
9	15.10	75.00	2.90	1.48	1.15	3.90	2.34
10	15.10	75.00	2.90	1.20	0.59	9.90	4.53
16	37.40	54.00	1.90	1.21	0.75	0.76	-0.10
19	42.90	46.00	1.60	1.10	0.55	0.76	-0.20
20	40.40	50.00	1.80	0.74	-1.14	0.55	-0.60
21	44.00	44.00	1.60	0.84	-0.86	1.05	0.29
22	48.00	36.00	1.40	1.34	2.07	1.25	0.56
23	37.40	54.00	1.90	0.72	-0.96	0.39	-0.83
24	46.10	40.00	1.50	1.36	2.03	0.98	0.17
25	42.30	47.00	1.60	0.97	-0.08	0.73	-0.24
26	47.50	37.00	1.50	1.04	0.31	0.75	-0.17
27	48.40	35.00	1.40	1.19	1.26	1.68	1.03
28	48.00	36.00	1.40	1.30	1.87	1.23	0.54
29	48.90	34.00	1.40	0.91	-0.57	0.63	-0.31
30	50.20	31.00	1.40	0.98	-0.09	0.83	0.07
31	55.30	20.00	1.50	1.03	0.24	0.85	-0.04
32	51.10	29.00	1.40	0.69	-2.42	0.42	-0.60
33	52.00	27.00	1.40	0.96	-0.20	0.64	-0.28
34	61.00	10.00	1.80	1.04	0.24	0.51	-0.83
35	60.30	11.00	1.80	1.21	0.97	0.64	-0.52
36	59.00	13.00	1.70	0.74	-1.38	0.37	-1.24
37	62.60	8.00	2.00	1.03	0.20	0.98	0.17
38	70.60	2.00	3.50	0.90	0.04	0.95	0.26
39	61.00	10.00	1.80	0.70	-1.40	0.28	-1.57
40	63.40	7.00	2.10	0.99	0.05	0.36	-1.10
41	70.60	2.00	3.50	1.06	0.29	0.44	-0.44
42	66.90	4.00	2.60	0.75	-0.56	0.19	-1.31
44	74.10	1.00	4.80	0.76	0.02	0.07	-1.34
45	74.10	1.00	4.80	0.76	0.02	0.07	-1.34

Note. Mnsq = mean squares; zstd = standardized values.

persons available for the regression analysis. Therefore, a decision was made to retain all the persons for the main analyses.

Table 10. *Rasch Analysis of Vocabulary: Item Statistics*

Item	Measure	Score	Error	INFIT mnsq	INFIT zstd	OUTFIT mnsq	OUTFIT zstd
4	31.50	68.00	1.30	0.80	-1.19	0.75	-1.12
7	30.00	72.00	1.30	0.70	-1.76	0.53	-2.30
9	40.00	41.00	1.20	1.29	2.13	1.35	1.79
11	52.50	11.00	1.80	1.02	0.17	0.73	-0.21
12	40.70	39.00	1.20	0.97	-0.22	0.95	-0.18
13	38.90	45.00	1.20	0.84	-1.27	0.78	-1.29
14	44.40	28.00	1.30	1.03	0.28	0.88	-0.36
15	34.90	58.00	1.20	1.07	0.53	1.14	0.78
16	40.30	40.00	1.20	0.96	-0.24	0.92	-0.39
17	42.30	34.00	1.20	1.14	1.05	1.11	0.57
20	47.50	20.00	1.40	1.18	1.01	1.13	0.44
22	59.90	4.00	2.70	0.87	-0.21	0.20	-0.73
23	50.00	15.00	1.60	1.04	0.27	0.66	-0.59
24	53.30	10.00	1.90	1.32	1.10	0.92	0.13
25	57.10	6.00	2.30	1.32	0.94	0.50	-0.31
26	61.70	3.00	3.00	1.17	0.51	0.39	-0.29
27	54.10	9.00	2.00	1.11	0.46	0.53	-0.51
28	67.70	1.00	4.80	1.05	0.35	0.17	-0.63
31	59.90	4.00	2.70	0.82	-0.34	0.20	-0.73
33	64.00	2.00	3.60	0.97	0.13	0.28	-0.44
34	61.70	3.00	3.00	0.72	-0.53	0.12	-0.90
40	67.70	1.00	4.80	0.81	0.05	0.07	-0.98

Note. Mnsq = mean squares; zstd = standardized values.

Principal Components Analysis of Residuals

A Rasch principal components analysis (PCA) of standardized residuals was performed to assess the dimensionality of each variable. PCA identifies the presence of multidimensionality by examining the characteristics shared in common among item residuals. In this analysis, each item is modeled to contribute one unit of randomness and there are as many residual variance units as there are items (Linacre, 2006, p. 243). When all information in the data is explained by the latent variable (i.e., when unidimensionality is present), residual values represent random noise and, when standardized, follow a normal distribution (Smith, 2004, p. 582).

Interpretation of the PCA results depends in part on the choice of the critical values for the eigenvalue and for the percentage of explained variance. Linacre (2006) provides a useful rule of thumb for the percentage of total variance accounted for by the measures for the construct to be considered unidimensional: Variance explained by measures should be more than 60% and unexplained residual variance explained by the first factor should be less than 5%. He further notes that in practice, a secondary dimension must have the strength of at least five items before these items can be considered a separate instrument (p. 289). Additionally, Graves and Weinstein (2004) recommend that for a construct to be considered unidimensional, the Rasch dimension should explain at least three times the variance explained by the first factor resulting from the PCA of residuals. Consequently in this study, a test was considered to be approximately

unidimensional if (a) it explained 60% or more of the total variance; (b) the unexplained variance explained by the first factor was less than 5%; (c) the eigenvalue for the first factor was less than 5; and (d) the Rasch dimension explained at least three times the variance explained by the first factor resulting from the PCA of residuals.

Table 11 shows the Rasch PCA results for the seven latent variables used in the study. All the variables except one conform to the criteria of unidimensionality specified above. The only exception is the DACC variable, for which the measurement dimension explains only 50.9% of the total variance. However, all the other criteria of unidimensionality are met: The unexplained variance explained by the first factor resulting from the PCA of residuals is 2.6%, the eigenvalue for the first factor is less than 5, and the Rasch dimension explains almost 20 times the variance explained by the first factor. Given that more than half of the variance in this variable was explained by the model, that the first factor explained only a very small percentage of the total variance, and that the Rasch dimension was much bigger than the identified sub-dimension, this sub-dimension was not expected to have much impact on person measurement. Thus, this variable was considered roughly unidimensional.

It should be noted that in each of the variables, a certain amount of unexplained variance (between .1 and 49.1%) remained after the Rasch dimension was extracted. This, however, is common for “broad” dimensions (Linacre, 2006, p. 312) such as those operationalized in this study. Further, as Linacre has pointed

out, multidimensionality is always present to some extent and the vital question is not whether the data are perfectly unidimensional but rather, whether the lack of unidimensionality in the data is sufficiently large to threaten the validity of the results. According to Linacre, multidimensionality only becomes a concern "when there are response patterns in the data indicating that the data represent two or more dimensions so disparate that it is no longer clear what latent dimension the Rasch dimension operationalizes" (Linacre, 1998, p. 268). This is clearly not the case in these data. Consequently, although a secondary dimension is somewhat noticeable in most of the variables used in this study, the principal components analysis indicates that the variables can be considered fundamentally unidimensional.

Table 11. *Standardized Residual Variance for Cognitive Ability and Decoding Measures (in Eigenvalue Units)*

Variable	Explained variance		Unexplained variance	Unexplained variance explained by 1st factor
	Empirical	Modeled		
PHAW	93.4 (62.9%)	62.2%	55.0 (37.1%)	4.3 (2.9%)
STM	539.0 (95.7%)	94.9%	24.0 (4.3%)	2.4 (.4%)
DACC	91.2 (50.9%)	49.4%	88.0 (49.1%)	4.6 (2.6%)
LISCOM	43.5 (74.3%)	73.6%	15.0 (25.7%)	2.0 (3.4%)
READCOM	54.3 (79.5%)	80.2%	14.0 (20.5%)	1.9 (2.8%)
VOC	548.6 (96.1%)	94.1%	22.0 (3.9%)	2.9 (.5%)
IQ	23186.9 (99.9%)	99.9%	30.0 (.1%)	3.2 (.0%)

Note. PHAW = phonological awareness, STM = short-term memory, DACC = decoding accuracy, LISCOM = listening comprehension, READCOM = reading comprehension, VOC = vocabulary, IQ = nonverbal ability.

In summary, the cumulative evidence from the evaluation of item fit and the PCA of standardized residuals indicates that each of the instruments used in this study measures a single construct. The decision to treat each of the tests as one variable is therefore justified.

Multivariate Analyses

Evaluation of Assumptions

As a result of variable construction, ten variables were available for major analyses: Phonological Awareness, Short-Term Memory, Decoding Accuracy, Listening Comprehension, Reading Comprehension, Vocabulary, Nonverbal Ability, RAN, RAN Rate, and Decoding Rate. The first seven were based on Rasch logits of person ability; the other three were based on time data. Prior to the analyses, these variables were examined through SPSS (Version 10.0) programs for accuracy of data entry, missing values, and the fit between their distributions and the assumptions of multivariate analysis. Three SPSS programs were used for this purpose: SPSS FREQUENCIES, SPSS REGRESSION, and SPSS DISCRIMINANT.

Although the original data contained several missing values for a small number of children, the logits of ability for these children were estimated through Rasch analysis. There were thus no missing values in the data. Prior to the multiple regression analyses, the data were assessed for normality, linearity, homoscedasticity of residuals, multicollinearity or singularity, and the presence of

outliers and suppressor variables. All the variables had relatively normally shaped histograms and satisfactory skewness and kurtosis values (see Table 12). Normal probability plots and residuals scatterplots were also examined for all the variables and no gross departures from normality, linearity, and homoscedasticity were detected (see Appendices Q-S for the scatterplots). Additionally, a White test was conducted to confirm the absence of heteroscedasticity in the regression models. This test tests the null hypothesis that the variance of the residuals is homogeneous. The results indicated that there was no heteroscedasticity in the models (χ^2 (5) = 3.94, p = 0.5 for the model with Decoding Accuracy as the dependent variable; χ^2 (14) = 12.95, p = 0.5 for the model with Decoding Rate as the dependent variable; χ^2 (14) = 14.95, p = 0.4 for the model with Reading Comprehension as the dependent variable). An evaluation of tolerance and VIF values revealed no multicollinearity problems (tolerance > .20; VIF < 4.0).

Two univariate outliers were detected through visual inspection of the boxplot for each variable. One had an extremely high score on DACC (78.43 CHIPS, z = 3.5), and the other had extreme scores on STM and READCOM (STM, 75.90 CHIPS, z = 3.3; READCOM, 27.95 CHIPS, z = -4.8). The first outlier turned out to be a high-ability person who performed consistently near the top of the distribution on all the tests. This person was considered to be part of the target population and was retained for all analyses. To reduce this person's influence on the regression model, his score was recoded to be one unit greater than the score of

the next extreme person in the distribution (new score = 70.61 CHIPS). This is an acceptable strategy for dealing with such cases (Tabachnick & Fidell, 2001).

The other outlier had numerous missing values on the STM and READCOM tests and her scores were estimated through the Rasch analysis. Unfortunately this estimation resulted in distortion. This was an average-ability person who only did the nonword repetition part of the STM test and the first three (easy) questions on the READCOM test. On the nonword repetition part of STM she unexpectedly scored much higher than average, whereas on READCOM, she unexpectedly missed two easy questions. Based on her average performance on most of the other tests, this person was unlikely to score much higher than average on the more difficult part of the STM test (i.e., digit repetition) or much lower than average on the READCOM test. I therefore replaced her values on these tests with the sample's mean (51.74 CHIPS for STM; 53.54 CHIPS for READCOM).

The data were also examined for multivariate outliers, or cases with an unusual combination of scores on two or more variables. Three influence statistics were used: Mahalanobis distance, Cook's D, and leverage. The criteria for multivariate outliers were Mahalanobis distance greater than χ^2 (4) = 18.47 ($p <$ .001) (Tabachnick & Fidell, 2001, p. 93), Cook's D greater than 1.0 (Cohen, Cohen, West, & Aiken, 2003), and leverage values greater than .18 (for a model with four predictors, Field, 2005). No cases exceeding the cut-off values were found in the data. The regression output was also examined for suppressor variables and for outliers in the solution using the statistical criterion of standardized

residuals in excess of +/-3.3 (Tabachnick & Fidell, 2001, p. 122). No suppressor variables or outliers were found.

Additionally, the data were evaluated for the assumptions of discriminant function analysis. The sample was split into two groups, a group of good readers (n = 38) and a group of poor readers (n = 41), using Decoding Rate as the criterion for classification, and assumptions were evaluated separately for each group. One univariate outlier was identified in the poor reader group (DRATE z < -.3.5); this case was deleted. To equalize the number of cases per group, two cases were randomly deleted from the poor reader group, which left 38 cases per group for analyses. An evaluation of assumptions of linearity, normality, homogeneity of variance-covariance matrices, and multicollinearity revealed no threat to multivariate analysis; there were no multivariate outliers in either group.

This procedure was repeated with more extreme reader groups. For this purpose, the sample was split into two groups separated from each other by a buffer zone of 1 standard deviation (0.5 *SD* above and below the mean) on DRATE. Children with z-scores above 0.5 on DRATE were classified as good readers whereas children with z-scores below -0.5 on DRATE were classified as poor readers. Children falling within the buffer zone were excluded from further analyses. This procedure produced 24 poor readers and 21 good readers. Two cases in the poor reader group were identified as univariate outliers (DRATE z < -3.3), and two cases in the good reader group were identified as multivariate outliers at p < .001. These cases were deleted. To equalize the number of cases per group, three

cases were randomly deleted from the good reader group; this left 19 cases per group available for analyses. The evaluation of the assumptions of multivariate normality, linearity, homogeneity of variance-covariance matrices, and multicollinearity was satisfactory.

Ratio of Cases to Predictor Variables

An important question in conducting multiple regression analysis is what constitutes an adequate sample size. Several rules of thumb have been proposed. For example, Miller and Kunce (1973) recommended a subject-to-variable ratio of 10 to 1; Tabachnick and Fidell (1989) provided "a bare minimum requirement" (p. 128) of five cases for each IV. However, several statisticians have expressed concern over the adequacy of such rules of thumb, calling them unsatisfactory (Green, 1991, p. 501) or even "mathematically incorrect" (Milton, 1986, p. 117). The main problem with such rules is that they do not take into account effect size. Yet, as Milton has argued, the determination of sample size is "not simply a function of the statistical technique one is using" because "there are real parameters in the population" (p. 113). Calculating the required sample size, therefore, necessitates making an a priori assessment of the population parameters that are being estimated. Furthermore, a sample must be adequate to enable both testing hypotheses with adequate power and estimating regression coefficients with adequate accuracy, and meeting these two goals may require different sample sizes (Algina, Moulder, & Moser, 2002).

Two different methods were used in this study to calculate the minimum sample size needed to ensure both adequate power and adequate estimation accuracy. To calculate the sample size needed for prediction accuracy, a formula developed by Sawyer (1982) was used, which is based on the mean absolute error of prediction as a measure of prediction accuracy. In this formula, N is a function of the number of predictors p and an inflation factor K, which is a measure of prediction accuracy. For $K = 1.05$, $N = 10.8p + 11.8$. Thus, for a model with four predictors, if the expected mean absolute error of prediction is to be $K = 1.05$ times that associated with a prediction equation based on known population values of the four regression coefficients, the minimum sample size equals 55.

To calculate the sample size needed to obtain adequate power, a method proposed by Green (1991) was used. This method requires specifying three values: alpha, or the probability of incorrectly rejecting the null hypothesis, power, or one minus the probability of not rejecting a false null hypothesis, and effect size, or the degree to which the criterion variable is related to the predictor variables in the population. The following formulas were used to calculate the required sample size with $\alpha = .05$ and a power of .80, a value described by Cohen (1988, 1992) as appropriate for behavioral research: $N \geq (8 / f^2) + (m - 1)$ for testing partial correlation coefficients and $N \geq L / f^2$ for evaluating the multiple correlation (where $L = 6.4 + 1.65m - .05m^2$, or 12.2 for $m = 4$, and $f^2 = R^2 / (1-R^2)$. Cohen's (1988) recommended values of $R^2 = .13$ and $f^2 = .15$ were used for a medium effect size.

From these formulas, the minimum sample size needed to test the multiple correlation with $m = 4$ was $N \geq (8 / .15) + (4 - 1) = 56$, and that needed to test individual predictors was $N \geq 12.2 / .15 = 81$. In this study, even greater effect sizes were expected because of the careful selection of potential predictors and exclusion of ambiguous ones (e.g., speech rate, visual processing). Therefore, $N = 79$ was considered appropriate.

Analysis Procedures

Standard multiple regression analyses were performed to answer Research Questions 1 and 2. This approach was chosen because the purpose was to assess the overall relationship between cognitive abilities and reading rather than to test an explicit hypothesis. Sequential regression analyses were performed to answer Research Question 4. Sequential regression was chosen in order to test a specific model; the choice of the predictors and the order of their entry into the regression equation were guided by theory.

Two sets of direct discriminant function analyses were performed to answer Research Question 5: In one, the whole sample was split into two groups, a group of good readers and a group of poor readers; in the other, a 1-*SD* buffer zone was used to define more extreme groups of good and poor readers. A classification function was then derived for the original sample. All analyses were performed using SPSS (Version 10.0).

Logits of ability (converted to CHIPS) were used in all statistical analyses for the following variables: Phonological Awareness, Short-Term Memory, Decoding Accuracy, Listening Comprehension, Reading Comprehension, Vocabulary, and Nonverbal Ability. For RAN, the scores were the speed (in seconds) of naming all 100 stimuli; for RAN Rate, the scores were the number of stimuli on both RAN tests named in one minute; and for Decoding Rate, the scores were the number of words and nonwords read correctly in one minute.

To answer Research Question 6, the children's *z*-scores on DACC and DRATE were examined to identify extremely poor readers in the sample. Following Grigorenko (2003), *extremely poor reading* was operationalized as a score on either DACC or DRATE (or both) that was at least 2.5 *SD*s below the mean. One child matched this definition. This was the same girl who was excluded from discriminant function analyses as a univariate outlier (see above). The child's cognitive, linguistic, and reading profile was then examined and compared with that of the good and poor readers to identify relative strengths and weaknesses.

CHAPTER 5

RESULTS

Descriptive Statistics

Table 12 shows the mean, the standard error of the mean, the standard deviation, the skewness, the standard error of skewness, the kurtosis, and the standard error of kurtosis for each cognitive, linguistic, and decoding measure used in this study. Overall, the statistics support the psychometric adequacy of the tests used in this study for the present sample. Performance on all tests was free from floor or ceiling effects, and all variables are relatively normally distributed, indicating that the mean provides an accurate estimate of central tendency. Two variables, Decoding Rate and IQ, have relatively large standard deviations, indicating that there is somewhat more variability in these measures than in the others.

The mean for Decoding Rate appears to be much lower than the norm of 30 words per minute that children in Russian public schools are expected to attain at the end of first grade (Alekseenko et al., 2003). This discrepancy, however, is likely a result of the differences in the testing instruments used. The Russian norm is based on the reading of a grade-appropriate passage containing many orthographically easy and visually familiar words embedded in a meaningful context. In contrast, the reading tests used in this study contained many visually unfamiliar, long, and morphologically and orthographically complex items, which

required letter-by-letter processing, and there was no contextual support. The mean and standard deviation values for word reading rate obtained in this study were much higher, $M = 25$; $SD = 10$, showing that at least some of the children had achieved the prescribed reading level.

Table 12. *Descriptive Statistics for Cognitive Ability and Decoding Measures* ($N = 79$)

Variable	*M*	*SE*	*SD*	Skewness	*SES*	Kurtosis	*SEK*
Decoding accuracy	60.1	.55	4.8	-.144	.271	.074	.535
Decoding rate	20.6	.79	7.0	.389	.271	.581	.535
Phonological awareness	56.3	.62	5.5	.071	.271	.125	.535
RAN	89.1	1.46	12.9	.406	.271	-.603	.535
RAN rate	68.7	1.10	9.8	.113	.271	-.779	.535
Short-term memory	51.4	.77	6.9	.552	.271	.278	.535
Listening comprehension	51.2	.38	3.4	-.337	.271	.380	.535
Reading comprehension	53.5	.49	4.4	.264	.271	-.063	.535
Vocabulary	27.4	.86	7.6	.623	.271	.496	.535
IQ	43.8	1.56	13.9	-.661	.271	-.570	.535

Note. Scores on all tests except RAN, RAN rate, and decoding rate are logits of ability (converted to CHIPS). RAN is the speed (in seconds) of naming 100 stimuli; RAN rate is the number of stimuli named in 60 seconds; decoding rate is the number of items read correctly in one minute. To facilitate comparisons with previous research, IQ scores were also converted into standard scores ($M = 100$; $SD = 15$). The range for the standardized IQ scores was 71-128.

Predictors of Decoding Accuracy

Research Question 1a asked about the size of the overall relationship between decoding accuracy and the set of phonological awareness, verbal short-term memory, rapid automatized naming, and IQ in beginning Russian readers.

Research Question 1b asked how much of the relationship was contributed uniquely by each predictor. To answer these questions, a standard multiple regression was performed with decoding accuracy as the criterion variable.

Table 13 shows the intercorrelations between decoding accuracy, phonological awareness, short-term memory, rapid automatized naming (RAN), and IQ. Note that superior performance on the RAN measure is indicated by lower scores and thus its correlation with decoding accuracy is negative. Only two variables correlate statistically significantly with decoding accuracy, phonological awareness and RAN; the other two variables, IQ and short-term memory, do not correlate with decoding accuracy at all. Because multiple regression builds on correlation and because the goal is to identify the smallest set of predictors that correlate strongly with the criterion variable, IQ and STM were excluded from the regression analyses.

Table 13. *Means, Standard Deviations, and Intercorrelations for Decoding Accuracy and Cognitive Predictor Variables* ($N = 79$)

Variable	*M*	*SD*	1	2	3	4
Decoding accuracy	60.1	4.8	.13	.29**	-.26*	.09
Predictor variable						
1. IQ	43.8	13.9	--	.41**	-.37**	.19
2. Phonological awareness	56.3	5.5		--	-.28*	.35**
3. RAN	89.1	12.9			--	-.16
4. Short-term memory	51.4	6.9				--

** $p < .01$. * $p < .05$.

A standard multiple regression was performed using decoding accuracy as the criterion variable and phonological awareness and RAN as predictor variables. Table 14 displays the unstandardized regression coefficients (*B*), the standard error of *B*, the standardized regression coefficients (β), the semipartial correlations (sr^2), and *R*, R^2, and adjusted R^2. *R* for regression was significantly different from zero, $F(2, 78) = 5.190$, $p = .008$. For the regression coefficient that differed significantly from zero, 95% confidence limits were calculated. The confidence limits were .016 to .405.

Table 14. *Standard Multiple Regression of Cognitive Variables on Decoding Accuracy* ($N = 79$)

Variable	*B*	*SEB*	*Beta*	sr^2 *(unique)*
Phonological awareness	.211*	.10	.24	.05
RAN	-.007	.04	-.19	

$R^2 = .12$. Adjusted $R^2 = .10$. $R = .35$**.

**$p < .01$. * $p < .05$.

Only one of the predictor variables, phonological awareness, contributed significantly to the prediction of decoding accuracy, with $t = 2.151$, $p = .035$. This variable contributed 5% of unique variance ($sr^2 = .05$) to the prediction. The two predictor variables in combination contributed another .07 in shared variability. Altogether, 12% (10% adjusted) of the variability was predicted by knowing scores on these two variables. Although the correlation between decoding accuracy and

RAN was significantly different from zero (r = -.26, p < .05), RAN did not contribute significantly to the regression. Apparently, the relationship between decoding accuracy and RAN is mediated by the relationship between decoding accuracy and phonological awareness.

Predictors of Decoding Rate

Research Question 2a asked about the size of the overall relationship between decoding rate and the set of phonological awareness, verbal short-term memory, rapid automatized naming (RAN) rate, and IQ in beginning Russian readers. Research Question 2b asked how much of the relationship was contributed uniquely by each predictor. To answer these questions, a standard multiple regression was performed with decoding rate as the criterion variable.

Table 15. *Means, Standard Deviations, and Intercorrelations for Decoding Rate and Cognitive Predictor Variables* (N = 79)

Variable	*M*	*SD*	1	2	3	4
Decoding rate	20.6	7.0	.21	.43**	.49**	.19
Predictor variable						
1. IQ	43.8	13.9	--	.41**	.36**	.19
2. Phonological awareness	56.3	5.5		--	.30**	.35**
3. RAN rate	68.7	9.8			--	.19
4. Short-term memory	51.4	6.9				--

**p < .001.

Table 15 shows the intercorrelations between decoding rate, phonological awareness, short-term memory, RAN rate, and IQ. Overall, the correlations between the predictor variables and decoding rate are higher than the correlations between these same predictors and decoding accuracy (see Table 13), although here too only two predictors correlate statistically significantly with the criterion variable: phonological awareness and RAN rate. To investigate the contribution of each predictor variable to the prediction of decoding rate, a standard multiple regression was performed between decoding rate as the criterion variable and phonological awareness, short-term memory, RAN rate, and IQ as predictor variables.

Table 16. *Standard Multiple Regression of Cognitive Variables on Decoding Rate* ($N = 79$)

Variable	*B*	*SEB*	*Beta*	sr^2 *(unique)*
Phonological awareness	.437**	.14	.34	.09
RAN rate	.298**	.08	.41	.14
IQ	-.004	.06	-.08	
Short-term memory	.000	.10	.00	

$R^2 = .33$. Adjusted $R^2 = .30$. $R = .58$**.

*$p < .05$. **$p < .01$.

Table 16 displays the unstandardized regression coefficients (*B*), the standard error of *B*, the standardized regression coefficients (β), the semipartial correlations (sr^2), and *R,* R^2, and adjusted R^2. *R* for regression was significantly

different from zero, F (4, 78) = 9.270, p = .000. For the regression coefficients that differed significantly from zero, 95% confidence limits were calculated. The confidence limits for phonological awareness were .159 to .715 and those for RAN rate were .149 to .446. Only two of the predictors contributed significantly to the prediction of decoding rate, phonological awareness (sr^2 = .09) and RAN rate (sr^2 = .14). The four predictor variables in combination contributed another 10% in shared variability. Altogether, 33% (30% adjusted) of the variability was predicted by knowing scores on these four predictor variables.

The Structure of Phonological Awareness

Research Question 3 asked whether rhyme and phoneme detection, phoneme segmentation, and phoneme deletion tasks tapped the same underlying ability in the Russian language. To answer this question, Rasch analysis was performed on the phonological awareness tests.

The item fit statistics for the whole phonological battery revealed eight misfitting items, or 5% of the test. All of the misfitting items required the detection of a sound unit (rhyme or phoneme). After the deletion of these items, a much better calibration was obtained, with only two items showing a slight misfit (outfit mean squares of 1.73 and 1.93). The fit statistics thus indicate that the phonological tasks used in this study tap the same underlying ability in the Russian language.

This is confirmed by the results of the PCA of standardized residuals (Table 11) showing that taken together, the tasks explain over 60% of the variance in the

phonological awareness variable. Although the results indicate the presence of a subdimension, it constitutes less than 3% of the variance. This does not mean that phonological awareness is a multidimensional concept; rather, the presence of a few off-dimension items indicates the "fuzzy," or broad nature of the concept of phonological awareness (Linacre, 2006). Thus, based on the cumulative evidence from fit statistics and PCA of residuals, rhyme and phoneme detection, phoneme segmentation, and phoneme deletion appear to tap a single underlying ability in the Russian language, which is approximately unidimensional.

Predictors of Reading Comprehension

Research Question 4 asked how accurately reading comprehension could be predicted (a) from IQ, (b) from listening comprehension and vocabulary over and above IQ, (c) from decoding rate over and above IQ, listening comprehension, and vocabulary, and (d) from the combination of IQ, listening comprehension, vocabulary, and decoding rate. To answer this question, a sequential multiple regression was performed using reading comprehension as the criterion variable.

Table 17 shows the intercorrelations between reading comprehension, IQ, listening comprehension, vocabulary, and decoding rate. The correlations between reading comprehension and all the predictor variables are statistically significant and moderate in magnitude. The intercorrelations between most of the predictor variables are also statistically significant, ranging from about .30 to about .60.

Table 17. *Means, Standard Deviations, and Intercorrelations for Reading Comprehension and Cognitive Predictor Variables* ($N = 79$)

Variable	*M*	*SD*	1	2	3	4
Reading comprehension	53.5	4.4	.54**	.54**	.52**	.37**
Predictor variable						
1. IQ	43.8	13.9	--	.48**	.61**	.21
2. Listening comprehension	51.2	3.4		--	.46**	.31**
3. Vocabulary	27.4	7.6			--	.33**
4. Decoding rate	20.5	7.0				--

**$p < .01$.

A sequential multiple regression was performed to determine how accurately reading comprehension could be predicted from the specified predictor variables. The predictors were entered in the following order: IQ was entered in step 1 as a control variable, vocabulary and listening comprehension were entered in step 2 in one block to determine if these abilities added significantly to the prediction of reading comprehension after the differences among the children in IQ were statistically eliminated, and decoding rate was entered in step 3 to determine if it added significantly to the prediction of reading comprehension after the differences among the children in IQ and linguistic ability were statistically eliminated.

Table 18 displays the unstandardized regression coefficients (B), the standard error of B, the standardized regression coefficients (β), the semipartial correlations (sr^2), and R, R^2, and adjusted R^2 after entry of all four predictor

variables. *R* was significantly different from zero at the end of each step. After step 3, with all predictors in the equation, $R = .67$, $F\ (4, 78) = 15.011$, $p = .000$.

Table 18. *Sequential Multiple Regression of Cognitive Variables on Reading Comprehension* ($N = 79$)

Variable	*B*	*SEB*	*Beta*	sr^2 *(incremental)*
Step 1				
IQ	.008	.04	.27	.29**
Step 2				
Vocabulary	.009	.13	.28	.13**
Listening comprehension	.362	.07	.17	
Step 3				
Decoding rate	.104	.06	.17	.02

R^2 =.45. Adjusted $R^2 = .42$. $R = .67$**.

**$p < .01$.

After step 1, with IQ in the equation, $R^2 = .29$, $F_{inc}\ (1, 77) = 32.131$, $p = .000$. Thus, IQ explained 29% of the variance in reading comprehension. After step 2, with vocabulary and listening comprehension added to the prediction of reading comprehension by IQ, $R^2 = .42$, $F_{inc}\ (2, 75) = 8.412$, $p = .001$. The addition of these measures to the equation with IQ resulted in a significant increment in R^2 and together, these measures explained 13% of variance over and above that explained by IQ. Within this block, however, the contribution of vocabulary was not statistically significant ($t = 1.429$, $p = .157$), although the correlation between reading comprehension and vocabulary was significantly different from zero ($r =$

.52 $p < .01$). Apparently, the relationship between reading comprehension and vocabulary is mediated by the relationship between reading comprehension and listening comprehension. After step 3, with decoding rate added to the prediction of reading comprehension by IQ and vocabulary and listening comprehension, R^2 = .45 (adjusted R^2 = .42), F_{inc} (1, 74) = 3.250, p = .075. The addition of decoding rate to the equation did not reliably improve R^2.

Cognitive and Linguistic Abilities of Good and Poor Readers

Research Question 5a asked to what degree there were differences between good and poor readers in the population on linear combinations of IQ, phonological awareness, listening comprehension, vocabulary, and reading comprehension. Research Question 5b asked if the children in the two groups could be correctly classified as good and poor readers based on their scores on the five predictor variables. To answer these questions, two sets of direct discriminant function analysis were performed with decoding rate as the grouping variable.

The sample was split into two groups, good readers and poor readers, using the participants' z-scores on decoding rate. Decoding rate was used to operationalize reading skill because it was hypothesized that this measure accurately captures the nature of the word recognition process: To be successful, readers need to be able to read both familiar and unfamiliar words accurately and quickly. A direct discriminant function analysis was performed using five cognitive variables as predictors of membership in the two groups. The predictors were

phonological awareness, reading comprehension, vocabulary, listening comprehension, and IQ. All analyses were performed using SPSS DISCRIMINANT.

Table 19. *Correlation of Predictor Variables with Discriminant Function (Function Structure Matrix) and Standardized Discriminant Function Coefficients* (n = 76)

Predictor variable	Correlations of predictors with discriminant function	Standardized discriminant function coefficients
Phonological awareness	.796	.566
Reading comprehension	.822	.587
Vocabulary	.681	.027
Listening comprehension	.539	.127
IQ	.595	-.050

Canonical R = .42; eigenvalue = .21

A discriminant function was calculated to determine whether the five predictors could predict decoding ability. Wilks's lambda was significant, Λ = .86, χ^2 (5, n = 76) = 13.725, p = .017, indicating that overall, the predictors differentiated between the two reader groups. The loading matrix of correlations between the predictors and the discriminant function is shown in Table 19. It indicates that the best predictors for distinguishing between the good and poor readers in the sample are reading comprehension, phonological awareness, and vocabulary.

Table 20. *Means and Standard Deviations of Predictor Variables as a Function of Decoding Ability* (n = 76)

Variable	Good readers (n = 38)		Poor readers (n = 38)	
	M	*SD*	*M*	*SD*
Decoding rate	26.23	4.97	15.69	3.30
Predictor variable				
1. Phonological awareness	58.08	5.30	54.28	5.20
2. Reading comprehension	55.00	4.26	51.95	3.92
3. Vocabulary	29.30	7.25	24.90	6.98
4. Listening comprehension	52.10	2.98	50.44	3.74
5. IQ	46.90	12.75	39.63	14.15

Note. All group means differ significantly at $p < .05$.

The means on the discriminant function are consistent with this interpretation (Table 20). Good readers have better reading comprehension ability (M = 55.00, SD = 4.26), better phonological skills (M = 58.08, SD = 5.30), and better vocabulary (M = 29.30, SD = 7.24) than poor readers (M = 51.95, SD = 3.92 on reading comprehension; M = 54.28, SD = 5.20 on phonological awareness; M = 24.90, SD = 6.98 on vocabulary). The groups also differ reliably on the other two measures, IQ (M = 46.90, SD = 12.75 for the good reader group; M = 39.63, SD = 14.15 for the poor reader group) and listening comprehension (M = 52.10, SD = 2.98 for the good reader group; M = 50.44, SD = 3.74 for the poor reader group), although the magnitude of these differences is smaller. The eigenvalue for the discriminant function is .21 and the canonical correlation is .42, indicating that about 18% of the variability of scores for the discriminant function is accounted for by the differences between the groups.

Table 21 shows pooled within-group correlations among the five predictors. All ten correlations would show statistical significance at $\alpha = .01$ if tested individually. All correlations are positive and most are in the .40-.58 range, indicating that the children who performed well on one of the measures tended to perform well on all the other measures, and that those who performed poorly on one of the measures tended to perform poorly on all the other measures. This is particularly true of the relationship between vocabulary and IQ (r (74) = .58, $p <$ .01) and vocabulary and phonological awareness (r (74) = .56, $p < .01$). This pattern indicates that children with strengths in vocabulary are likely to have higher nonverbal intelligence and better phonological awareness skills.

Table 21. *Pooled Within-Group Correlations Among Predictors* (n = 76)

Predictor variable	1	2	3	4	5
1. Phonological awareness	--	.33	.56	.39	.33
2. Reading comprehension		--	.48	.51	.49
3. Vocabulary			--	.45	.58
4. Listening comprehension				--	.46
5. IQ					--

A classification function was derived for the original sample, with 64.5% of the children in the sample classified correctly. In order to take into account chance agreements, a kappa coefficient was computed. This coefficient assesses the accuracy of classification, with 1 indicating a perfect accuracy and 0 indicating chance-level prediction (Green & Salkind, 2005). The obtained value was .30,

indicating moderate accuracy in prediction. The stability of the classification procedure was assessed with a cross-validation run. The leave-one-out technique was used, in which the classification function is derived on the basis of all cases except one, and then the left-out case is classified. This procedure is repeated until all cases have been left out once and classified. Based on this procedure, 58% of all cases were classified correctly.

Table 22. *Correlation of Predictor Variables with Discriminant Function (Function Structure Matrix) and Standardized Discriminant Function Coefficients* (n = 38)

Predictor variable	Correlations of predictors with discriminant function	Standardized discriminant function coefficients
Phonological awareness	.898	1.076
Reading comprehension	.619	.503
Vocabulary	.638	-.272
Listening comprehension	.375	-.469
IQ	.478	.152

Canonical R = .65; eigenvalue = .72.

To determine whether the same pattern of cognitive abilities characterized more extreme groups of good and poor readers, a second set of analyses were conducted with the groups of good and poor readers separated from each other by a buffer zone of 1 standard deviation (0.5 *SD* above and below the mean) on decoding rate. With univariate and multivariate outliers deleted, there were 19 cases per group.

A direct discriminant function analysis was performed using phonological awareness, reading comprehension, vocabulary, listening comprehension, and IQ as predictors of membership in the good and poor reader groups. A discriminant function was calculated to determine whether the five predictors could predict decoding ability. Wilks's lambda was significant, $\Lambda = .58$, χ^2 (5, $n = 38$) = 18.175, $p = .003$, indicating that overall, the predictors differentiated between the two reader groups. The loading matrix of correlations between the predictors and the discriminant function is shown in Table 22. It indicates that the best predictors for distinguishing between the good and poor readers in the sample are again phonological awareness, vocabulary, and reading comprehension.

Table 23. *Means and Standard Deviations of Predictor Variables as a Function of Decoding Ability* ($n = 38$)

Variable	Good readers ($n = 19$)		Poor readers ($n = 19$)	
	M	*SD*	*M*	*SD*
Decoding rate	29.79	4.74	13.85	2.26
Predictor variable				
1. Phonological awareness	60.53	5.19	52.50	5.63
2. Reading comprehension	55.29	3.23	51.34	4.39
3. Vocabulary	31.20	7.32	22.94	8.31
4. Listening comprehension	52.13	3.26	49.90	3.91
5. IQ	48.62	10.37	37.99	15.94

Note. All group means differ significantly at $p < .05$.

The means on the discriminant function are consistent with this interpretation (Table 23). Good readers have better phonological awareness skills

(*M* = 60.53, *SD* = 5.19), better vocabulary (*M* = 31.20, *SD* = 7.32), and better reading comprehension ability (*M* = 55.29, *SD* = 3.22) than poor readers (*M* = 52.50, *SD* = 5.63 on phonological awareness; *M* = 22.94, *SD* = 8.31 on vocabulary, and *M* = 51.34, *SD* = 4.39 on reading comprehension). The groups also differ reliably on IQ, with *M* = 48.62, *SD* = 10.37 for the good reader group and *M* = 37.99, *SD* = 15.94 for the poor reader group, although the magnitude of these differences is smaller. The eigenvalue for the discriminant function is .72 and the canonical correlation is .65, indicating that approximately 42% of variability of scores for the discriminant function is accounted for by the differences between the groups.

Table 24. *Pooled Within-Group Correlations Among Predictors* (*n* = 38)

Predictor variable	1	2	3	4	5
1. Phonological awareness	--	.43	.68	.58	.42
2. Reading comprehension		--	.64	.55	.58
3. Vocabulary			--	.49	.61
4. Listening comprehension				--	.53
5. IQ					--

Table 24 shows pooled within-group correlations among the five predictors. All ten correlations would show statistical significance at $\alpha = .01$ if tested individually. All correlations are positive and most are in the .50-.68 range. The pattern is very similar to that observed when the whole sample was classified into two groups, but the magnitude of the correlations appears to be larger. This is

particularly true of the relationship between vocabulary and IQ (r (38) = .61, p < .01), vocabulary and reading comprehension ((r (38) = .64, p < .01), and vocabulary and phonological awareness (r (38) = .68, p < .01). This pattern indicates that children with strengths in vocabulary are likely to have higher nonverbal intelligence, better reading comprehension ability, and stronger phonological awareness skills.

A classification function was derived for the original sample, with 81.6% of the children classified correctly. To take into account chance agreements, a kappa coefficient was computed. The obtained value was .63, indicating a high degree of accuracy in prediction. The stability of the classification procedure was assessed with a cross-validation run using the leave-one-out technique. Based on this procedure, 76.3% of all cases were classified correctly.

Cognitive and Linguistic Profile of the Poorest Reader

Research Question 6 asked about the cognitive, linguistic, and reading profile of the poorest readers in the sample. To answer this question, the children's z-scores on decoding accuracy (DACC) and decoding rate (DRATE) were examined to identify extremely poor readers in the sample. For this purpose, *extremely poor reading* was operationalized as a score on either DACC or DRATE (or both) that was at least 2.5 *SD*s below the mean (Grigorenko, 2003). Only one child, a girl (hereinafter referred to as NY), matched this definition. Her reading profile was then examined to identify areas of particular difficulty and her

performance on the cognitive and linguistic measures was compared to that of the children who scored above the mean on DRATE (n = 38) and those who scored .5 *SD* below the mean on DRATE (n = 19).

Table 25. *NY's Reading Performance in Comparison to That of the Rest of the Children*

Reading test (number of items)	Speed (NY)		Accuracy (%)	
	Time (*s*)	*z*-score	NY	RC
Word reading (50)	229	2.84	82	89
List 2 (morphologically simple, 10)	41	4.4	100	93
List 4 (morphologically complex, 10)	49	1.5	70	77
Nonword reading (48)	765	7.53	0	79
List 1 (visually dissimilar, 24)	470	7.71	0	79
List 2 (visually similar, 24)	295	6.88	0	79

Note. RC = rest of the children excluding NY (n = 78). Accuracy is the percentage of items read correctly.

NY was severely impaired on both decoding rate (absolute score = 2.47; z = -2.57) and decoding accuracy (Rasch person ability score = 47.02; z = -2.71). An examination of her reading performance on each test (Table 25) revealed several patterns. On word reading, NY's main impairment was in speed, rather than in accuracy. In fact, her overall reading accuracy was comparable to that of the rest of the children (NY: 82%; rest of the children: 89%) and her accuracy on the morphologically simple, one- and two-syllable words was even better than that of the rest of the sample. She was less accurate on the morphologically complex, three- and four-syllable words, but still managed to read 70% of them correctly.

The few mistakes that she made on both types of word were minor and included placement of stress on the wrong syllable and incorrect pronunciation of word endings (e.g., reading *prevratit'* [to turn] as *prevratil* [(he) turned].

Table 26. *NY's Performance on Cognitive and Linguistic Tests in Comparison to That of Good and Poor Readers*

Variable	NY	Poor readers (n = 19)		Good readers (n = 38)	
		M	*SD*	*M*	*SD*
Phonological awareness	58.32	52.50	5.63	58.08	5.30
Short-term memory	50.29	51.19	6.86	51.58	7.24
RAN rate	54.05	65.62	8.54	72.99	9.13
IQ	51.26	37.99	15.94	46.90	12.75
Listening comprehension	47.63	49.90	3.91	52.10	2.98
Reading comprehension	51.49	51.34	4.39	55.00	4.26
Vocabulary	27.32	22.94	8.31	29.30	7.25

Note. Scores on all tests except RAN rate are logits of ability (converted to CHIPS); RAN rate is number of stimuli named in 60 seconds.

In contrast, on nonword reading, NY was severely impaired on both speed and accuracy: She took almost 13 minutes to read 48 items, scoring 7 *SD* below the mean, and she failed to read a single nonword correctly. Most errors that she made were substitutions of the target items with real words that were visually similar to the target items or that contained some of the letters from the target items. For example, she read *bashai* as *bol'sho*i [big], *stur* as *stul* [chair], and *yuskra* as *iskra* [spark]. She was equally inaccurate on visually similar and visually dissimilar nonwords, although she was much slower on the dissimilar items than on the

similar ones. No other child in the sample was as impaired as NY on nonword reading.

NY's performance on the cognitive and linguistic tests is shown in Table 26. As shown in the table, on the phonological awareness, short-term memory, IQ, and vocabulary tests, NY performed close to the mean of the good readers in the sample whereas on the reading and listening comprehension tests she performed close to the mean of the poor readers. The variable that most clearly separated NY from both the good and poor readers was RAN rate: She was more than 2 *SD* below the mean of the good readers and more than 1 *SD* below the mean of the poor readers.

Can NY be classified as dyslexic? In reading research, it is common to operationalize dyslexia on the basis of a discrepancy between a person's actual reading score and that predicted based on their IQ (Fawcett, 2002). An arbitrary cut-off point of 1.5 or 2 standard deviations below the expected reading age is typically used to identify dyslexics. In this case, however, both the reading and IQ tests were unstandardized and there is no reference point for an accurate description. It does appear that the discrepancy between NY's above average IQ and her extremely slow and laborious reading would provide a basis for diagnosing her as dyslexic. The fact that she had particular difficulty with nonwords appears to confirm this classification.

CHAPTER 6

DISCUSSION

Predictors of Decoding

Research Question 1 asked about the size of the overall relationship between decoding accuracy and the set of phonological awareness, verbal short-term memory (STM), rapid automatized naming (RAN), and IQ in beginning Russian readers as well as about the unique contribution of each predictor. Research Question 2 asked about the size of the overall relationship between decoding rate and the set of phonological awareness, verbal short-term memory, RAN rate, and IQ in beginning Russian readers as well as about the unique contribution of each predictor. Two sets of standard multiple regression analyses were performed, first with decoding accuracy, and then with decoding rate, as the dependent variable.

In the first set of analyses, because of the low correlations between decoding accuracy and IQ and decoding accuracy and short-term memory, only two variables, phonological awareness and RAN, were included in the regression model. Together, the two variables predicted about 10% of the variability in decoding accuracy but only one of the predictors, phonological awareness, contributed a small amount of unique variance (i.e., about 5%). The other variable, RAN, did not make an independent contribution to the prediction although its correlation with decoding accuracy was statistically significant. In the second set of

analyses, the four predictor variables predicted about 33% of the variability in decoding rate; however, only two of the predictors, phonological awareness and RAN rate, contributed unique variance, and the contribution of the latter was larger than the contribution of the former (14% vs. 9%).

Taken together, these results indicate that phonological awareness is a reliable predictor of both decoding accuracy and decoding rate in beginning Russian readers. However, in contrast to English, where phonological tasks have been found to account for a substantial portion of variance in reading scores (Muter & Snowling, 1998; Tunmer et al., 1988), in Russian, the relation of these tasks to both decoding accuracy and decoding rate appears to be relatively small. This result is consistent with findings obtained in other transparent orthographies, where phonological awareness has been shown to develop very rapidly after the beginning of reading instruction (e.g., Caravolas & Bruck, 1993; Wimmer, 1993) and where individual differences in phonological skill appear to be relevant to reading only for a short period of time (de Jong, 2003; Durgunoglu & Oney, 1999; Holopainen et al., 2001; Share, 2003). The children in the present study were tested at the end of first grade and the beginning of second grade, or after one full year of reading instruction, and they found the decoding test quite easy as revealed by the Rasch analysis of item difficulty. It therefore seems likely that most of the children in the sample already had sufficiently high levels of phonological awareness for accurate decoding.

Further, phonological awareness in this study was more strongly related to decoding rate than to decoding accuracy, accounting for almost twice as much variance in decoding rate than in decoding accuracy. This result can also be attributed to the relative consistency of the Russian orthography. Previous research with Swedish (Olofsson, 2003) has shown than in contrast to decoding accuracy, which in consistent orthographies develops rapidly after the onset of reading instruction, decoding speed continues to grow throughout the elementary school years and often does not reach its asymptotic level until the end of elementary school. In younger children, however, high decoding accuracy levels are obtained at the expense of decoding speed. Thus, decoding speed in transparent orthographies may reflect not only decoding fluency but also the extent to which individual differences in phonological awareness remain relevant to reading. Empirical support for this view comes from a longitudinal study of reading in Dutch, a transparent orthography (de Jong & van der Leij, 1999), which found that once children attained a high decoding rate, phonological awareness ceased to have an additional effect on reading.

Compared to phonological awareness, rapid naming accounted for substantially more variance in decoding rate. There are two possible explanations for this result. The first one draws on a model of reading development proposed by Manis et al. (1999), in which RAN tasks are conceptualized as a measure of "a unique aspect of reading that is not captured by [phonological awareness] tasks, the learning of arbitrary mappings between print and sound" (p. 152). According to this

view, RAN tasks are more relevant than phonological tasks to the very early stages of reading development, when children are learning the arbitrary associations between symbols and sounds. In contrast, phonological tasks are relevant to both early and later reading development because they enable the children to build and access phonological representations of words. This explanation, however, does not sit well with the fact that most of the children in the present study were proficient decoders capable of decoding even phonologically complex words.

An alternative explanation is to consider the effect of orthography on the relationship between reading and rapid naming. It has been argued that in languages with consistent orthography, beginning readers rely almost exclusively on letter-by-letter decoding (e.g., Goswami et al., 2003; Ziegler et al., 2001) and can read words by simply assembling sequentially the names of the sounds that the letters represent. Consequently, the main characteristic that distinguishes good from poor readers in such orthographies is the speed of assembling word pronunciations (Wimmer, 1993; Wimmer & Goswami, 1994), a process that has been hypothesized to depend primarily on the rapid retrieval of phonological codes from long-term memory (de Jong & van der Leij, 2002). At its core, this process is similar to rapid naming tasks in that both require "quick access to visual-verbal associations" (de Jong & Vrielink, 2004) and high processing speed for efficiency. This explanation is consistent with the observation of the special importance of rapid naming as a predictor of reading achievement in transparent orthographies (e.g., Holopainen et al., 2001; Kobayashi et al., 2005).

Consistent with this interpretation is the present finding that rapid naming was more strongly related to decoding rate than to decoding accuracy. The correlation between rapid naming and decoding rate was almost twice as large as the correlation between rapid naming and decoding accuracy. Moreover, in the model with decoding rate as the criterion, rapid naming accounted for 14% of unique variance whereas in the model with decoding accuracy as the criterion, rapid naming did not explain any unique variance and its correlation with decoding accuracy was fully mediated by phonological awareness.

The last observation appears to indicate a certain amount of overlap between rapid naming and phonological awareness, an observation strengthened by the finding of a statistically significant correlation between these two variables. Similar results have been reported for English (Manis et al., 1999; Swanson et al., 2003). Nevertheless, the present results clearly show that the two measures account for independent variance at least in one measure of reading, decoding rate. To this end, they replicate previous findings that phonological awareness and rapid naming account for distinct variance in reading acquisition and lend some support to the argument of Wolf and Bowers (1999) that performance on RAN measures and performance on phonological awareness measures are the results of independent processes.

Perhaps the most striking—and somewhat unexpected—result obtained in this study is the failure of verbal STM to account for any variance either in decoding accuracy or decoding rate. At first glance this finding appears to

contradict the research evidence obtained in English and some other languages showing that verbal STM is an important correlate of reading (for reviews, see Singleton, 2002; Wagner & Torgesen, 1987). Yet the previous findings of a positive association between verbal STM and reading should not be taken to mean that there is an unequivocal understanding in the research community of how, exactly, memory processes contribute to reading development or, even, which components of memory relate to which aspects of reading. At least three explanations can be offered for the apparent lack of association between STM and decoding observed in the present study.

One explanation rests on the developmental nature of the relationship between reading and various cognitive abilities. According to this developmental view of reading (e.g., Muter & Snowling, 1998), different cognitive skills are relevant to different stages of reading development. Muter and Snowling argued that in the early stages, access to phonological segments of words combined with letter knowledge suffices for creating primitive mappings between orthography and phonology. At later stages, when children establish fine mappings between print and sound, memory becomes more important for specifying word representations at the phonemic level. This theory has found some empirical support. For example, Gathercole, Baddeley, and Willis (1991) found no relationship between reading and STM measured by nonword repetition and digit span in English-speaking children in the first few months of school but by the second year of school, there was a statistically significant correlation between these skills. De Jong and van der Leij

(1999) also reported a changing relationship between STM and reading in Dutch children. However, this explanation seems unlikely given that the children in this study had already had a full year of reading instruction and could hardly be considered truly beginning readers.

Perhaps a more plausible explanation lies in the nature of the tasks that were used to measure verbal STM. Although both the digit span task and the nonword repetition task used in this study have been described as measures of an individual's ability to maintain verbal information in short-term memory, there is some evidence (e.g., Hansen & Bowey, 1994) that these tasks may be tapping somewhat different aspects of verbal memory because digit span relies on the use of rehearsal strategies (Baddeley, 1986), whereas nonword repetition relies on the ability to form transient phonological representations (Gathercole & Baddeley, 1989). In the present study, Rasch analysis conducted to evaluate the STM test showed that both tasks tapped the same construct. However, it is still possible that although both tasks tapped the same underlying ability (i.e., to maintain verbal information in short-term memory), their relationship to decoding was not the same by virtue of the different strategies that they required for successful performance.

It can be speculated, for example, that nonword repetition would be much more closely related to reading than digit span because it requires access to the phonological structure of language. The hypothesized relationship between nonword repetition and underlying phonological representations may explain the statistically significant correlation that was observed between STM and

phonological awareness in the present study. In contrast, digit span would be associated with reading much less strongly because, as has been argued previously (Engle, Cantor, & Carullo, 1992), rehearsal strategies may not be generalizable to cognitive tasks like reading. The inclusion of the digit span task in the STM instrument in the present study may have lowered the correlation between the STM variable and decoding.

An alternative explanation that also stems from the nature of memory tasks draws on a distinction between short-term memory and working memory (e.g., Swanson, 1994). The former is defined as the ability to retain information in short-term store, whereas the latter is defined as the ability to carry out additional processing operations while holding information in short-term store. Swanson has shown that it is working memory rather than STM that is closely related to reading. In this view, repeating digits in the same order as they are presented would be a poor measure of an individual's working memory because such repetition is automatic and requires an immediate response (Kaufman, 1994). In contrast, phonological awareness tasks that require the child to select a word that does not "belong" from an orally presented string of three or four words could be said to include a working memory component because they require the simultaneous storage and processing of verbal information. In the present study, three of the five phonological tasks were of the "odd-one-out" variety. It is therefore possible that the working memory component in these phonological tasks obscured the effect of STM on reading. The reliance of these tasks on memory processes may also

explain the moderate correlation obtained between the measures of STM and phonological awareness. It also suggests that these tasks should not be construed purely as a measure of phonological skill but rather as a measure sensitive to the integrity of verbal memory processes.

An inconsistent relationship between verbal memory and reading has been noted in several non-English languages. For example, in Arabic, (Mannai & Everatt, 2005), forward digit span failed to correlate with word reading accuracy both in first- and third-grade children. In Japanese (Kobayashi et al., 2005), nonword repetition did not predict either reading speed or reading accuracy in kindergarten and first-grade children. In Finnish (Holopainen et al., 2001), nonword repetition and digit span measures obtained in preschool failed to predict nonword reading speed at the end of second grade although nonword repetition was statistically significant in differentiating ordinary from late decoders whereas digit span was statistically significant in differentiating precocious from late decoders. In Dutch (de Jong & van der Leij, 1999), verbal memory entered in a regression equation after phonological awareness did not account for any additional variance; in contrast, both phonological awareness and rapid naming did account for additional variance when memory scores were statistically controlled. In Hungarian (Everatt et al., 2004), good and poor readers did not differ at all on STM measured by a memory span task although they differed significantly on nonword reading. This was in contrast to monolingual English and bilingual English-Tagalog children for whom memory span was a reliable predictor of reading achievement.

The inconsistency of the relationship between verbal STM and reading observed in non-English languages introduces the possibility of a differential effect of verbal memory processes on reading development in different orthographies. It can be speculated, for example, that in opaque orthographies like English, the inconsistency of the sound-letter mappings forces the beginning reader to rely heavily on verbal memory to assemble word pronunciations. For example, in order to read words such as *though* and *tough*, the beginning reader would have to simultaneously hold in memory the various pronunciations of the letters and letter combinations involved, compare them in search of the right pronunciation, select the right one, and reject the ones that are incorrect. The efficiency of this processing is likely to depend on verbal short-term memory capacity. In contrast, in transparent orthographies such as Russian, readers can assemble word pronunciations letter-by-letter by simply pronouncing the sounds the letters represent. Children familiar with letter-sound names and with some access to phonology can perform this task without overloading the memory system even in the presence of STM deficits.

Another unexpected result obtained in the present study concerns the role of IQ in predicting decoding ability: IQ was found to be a poor predictor of both decoding accuracy and decoding rate. This result appears to contradict those obtained in English, most notably by Naglieri and Ronning (2000), who found moderately high correlations in excess of .40 between word reading and IQ for a sample of 22,000 school children. However, in that study, word reading was

assessed indirectly, through a vocabulary subsection on a standardized achievement test, which required not only word recognition but also the knowledge of synonyms and antonyms and the ability to understand word meanings, an ability that has been shown to be closely associated with reading comprehension (e.g., Torgesen et al., 1997). Thus, the word reading test in that study was not a pure measure of decoding.

It should be stressed, however, that some evidence of a connection between IQ and decoding was found in the present study: In the discriminant function analyses, the groups of good and poor readers differed reliably on IQ (see below for further discussion). This finding appears to be in agreement with previous research in English showing that IQ is much more relevant to the prediction of decoding ability in poor readers than in typical readers (Tiu et al., 2003). The substantial correlation between IQ and phonological awareness may also point—although indirectly—to the importance of IQ to some aspects of decoding, i.e., those that rely heavily on phonological processing.

Alternatively, the correlation between IQ and phonological awareness can be taken to indicate an effect of IQ on general information processing. It has been argued (e.g., Frith, 2002), for example, that performance on virtually any test would be affected by the general ability factor, or "g," which is the measurement target for most IQ tests. In this sense, IQ would be relevant to phonological awareness to the extent that performance on phonological tests relies on the use of effective information processing strategies. This interpretation would also be

consistent with Hammill's (2004) observation that phonological awareness, like all other processing abilities, may be a sign of general developmental maturity and may simply represent a different aspect of general intelligence.

The Structure of Phonological Awareness in the Russian Language

Research Question 3 asked whether rhyme and phoneme detection, phoneme segmentation, and phoneme deletion tasks tapped the same underlying ability in the Russian language. To answer this question, Rasch analysis was conducted to determine if the phonological tasks used in the present study formed a single variable. Additionally, the tasks and individual items were examined to identify factors that determined task and item difficulty.

The Rasch fit statistics as well as the PCA of standardized residuals clearly indicate that the phonological tasks of rhyme and phoneme detection, phoneme segmentation, and phoneme deletion used in the present study tap the same underlying ability, which is approximately unidimensional. This finding supports the conclusion of recent studies conducted with large samples of English-speaking children (e.g., Anthony & Lonigan, 2004; Anthony et al., 2002) that phonological awareness is a "single cognitive ability that manifests behaviorally in a variety of skills" (Anthony & Francis, 2005, p. 256) throughout a child's development and extends this view to the Russian language. It appears that, just as in English, phonological awareness in Russian can be conceptualized as a single ability that

can be measured by tasks requiring different cognitive operations (e.g., deletion, identification, segmentation) and with items varying in linguistic complexity.

The conceptualization of phonological awareness as a unidimensional construct carries with it several important implications. First, the distinctions that have been drawn between phonological tasks based on the cognitive operation they require or the linguistic unit that is the focus of the task may not be as important as has been previously thought, at least for children who are learning to read a transparent orthography and who have mastered the basic rules of print-to-sound conversion.

Second, the finding that both rhyme and phoneme tasks represent the same underlying ability supports Anthony and Lonigan's (2004) argument that phonological awareness may influence reading acquisition "through multiple pathways" (p. 214), including reading by analogy and reading by grapheme-phoneme decoding. This finding may be particularly important in light of the recent debate over the role of large versus small sound units in predicting reading development. Some researchers (e.g., Goswami & Bryant, 1990) have argued for the special importance of rhyme awareness whereas others (e.g., Muter et al., 1997; Nation & Hulme, 1997) have insisted on the special role of phonemes in predicting reading. Although the present study did not compare directly the importance of these skills in predicting reading, the finding that measures of rhyme and phoneme awareness represent the same underlying ability suggests—as has been argued previously by Anthony and his colleagues (Anthony & Lonigan; Anthony et al.,

2002)—that the debate over the importance of certain phonological tasks over others may have proceeded from an unwarranted assumption of the existence of different types of phonological awareness.

For practical purposes, these findings indicate that phonological awareness can be accurately measured by a variety of tasks, including those that are appropriate for young children who have not yet begun learning to read and who have not yet acquired high levels of phonemic awareness. This makes it possible to identify children with potential reading problems early on, before they fail in literacy acquisition. For older children, assessment tools should include a range of tasks with differing cognitive requirements and linguistic features so that the full extent of their phonological awareness can be evaluated.

It is important to point out, however, that the conceptualization of phonological awareness as a unitary ability does not imply that this skill may not include cognitive components besides phonological sensitivity. As has been suggested above, phonological awareness appears to be closely linked to verbal short-term memory and general cognitive ability and therefore phonological tasks may tap aspects of cognitive functioning that go beyond phonology. Furthermore, the uncovered unitary structure of phonological awareness should not be taken to mean that all phonological tasks are equally useful in tapping phonological ability in all children. As the item-person map of phonological awareness (see Figure 2) clearly indicates, some tasks are more informative about a child's phonological skills than others. The usefulness of the particular tasks may also be influenced by

age as well as reading experience (Anthony & Lonigan, 2004; Schatschneider et al., 1999).

The results of the Rasch analysis also provide useful information regarding the difficulty of the tasks that comprised the phonological awareness battery in this study. Overall, the battery appears to be too easy for this sample, with the mean person ability being almost 1 *SD* above the mean item difficulty and with one-third of the test being more than 2 *SD* lower on the logit scale than the mean ability of the children. This means that the children in the sample, who had had about one year of reading instruction, must have developed sufficient amounts of phonological sensitivity needed to perform common phonological tasks. This result is well supported by the theoretical accounts of reading development in languages with transparent orthographies in which children typically demonstrate excellent performance on phonological tasks—sometimes at ceiling—by the end of first grade (see Goswami, 2002, for a review).

Furthermore, the Rasch analysis clearly demonstrates the developmental nature of phonological awareness in the Russian language, with some tasks being more difficult than others. Syllable structure in particular appears to play a role in determining the degree of difficulty of an item. For example, as shown in the item-person map of phonological awareness (see Figure 2), many CVC items are clustered toward the bottom of the test, whereas many items with consonant blends appear closer to the top; items requiring the processing or manipulation of single sounds appear to be easier than those that require the processing of sound clusters.

This is particularly true of the phoneme segmentation task, a task that required the precise identification and separation of individual phonemes.

However, contrary to previous reports from English-language research, no clear hierarchy of task difficulty could be identified based on either the cognitive demands of a task or the size of the linguistic unit that was its focus. Rhyme tasks were no easier than phoneme tasks and those that required sound identification were no easier than those that required sound manipulation. This result runs counter to the notion of a hierarchy of phonological task difficulty proposed in English in which tasks are distinguished by both the cognitive operation required and the size of the linguistic unit that must be analyzed or manipulated (Adams, 1990; Yopp, 1988). In the present study, task and item difficulty depended on many factors including the position of the target item, the syllabic structure of the word, and the number of letters and sounds in the target word, and it was not determined solely by the cognitive demands of the task or the size of the linguistic unit that was its focus.

On the other hand, the item-person map of phonological awareness (Figure 2) shows a relationship between item difficulty and the placement of the target word or sound in the task. Overall, items with the target word or sound placed in the final position were easier across the board than those with the target word or sound placed in the initial position, which in turn were easier than those with the target unit placed in the middle position. This hierarchy held, although with some exceptions, for all tasks and regardless of the linguistic features of the items. For

example, of the 21 easiest items, which were 2 *SD* below the mean person ability, 18 (i.e., six RD, four IO, three FO, and five PD items) required the detection or manipulation of sound units with the target unit placed in the final position. In contrast, of the 12 hardest items, ten (five RD and five FO) had the target unit in the middle or initial position.

This result suggests that items with the target word or sound placed in the final position are easier to process than other items, possibly because such items can be processed on-line, sequentially. For example, in the "odd-word-out" tasks (e.g., rhyme detection), words in a string with the odd one placed in the final position (e.g., *play, clay, block*) can be compared in the order in which they are presented and do not require "going back" mentally to the beginning of the string upon hearing all the words in order to compare them. In deletion tasks, items with the phoneme to be deleted being in the final position can also be processed sequentially and such tasks can be performed by simply repeating the presented word without its final bit. While they clearly rely on the child's ability to understand the notion of subsyllabic units as well as to segment, these tasks do not necessitate complex phonological processing such as the breaking up of a word into constituent phonemes, the extraction of a target unit with the simultaneous holding of the remaining phonemes in memory, and the re-assembly of the new word, which would be needed in tasks requiring the deletion of a phoneme from the middle of a word. The burden on working memory may also be low in such tasks.

The finding that task difficulty may be affected by the position of the target item has important implications for the design and administration of phonological tests as it suggests the need for researchers to consider the order in which words are presented in phonological judgment tasks because of the possibility that this order may affect performance. In their analysis of the usefulness of various phonological awareness tasks in English, Schatschneider et al. (1999) came to a similar conclusion after discovering, using the IRT model, that the difficulty parameters of the sound categorization task, a task that required the children to select, out of four words presented, the word that did not sound like the others, were highly dependent on where the target word was placed in the string of words. The authors recommended that researchers avoid using such tasks because of possible confounding effects.

While it may not be practical to exclude such tasks from a phonological battery, especially with younger children, the results of the present study indicate that in designing tasks to measure phonological awareness, researchers must take into account factors beyond the cognitive nature of the task or the size of the linguistic unit to be analyzed, such as the position of the target word or sound or the syllabic complexity of the target items. These factors are likely to influence performance on phonological tasks and their effect would be difficult to separate from the effect of the cognitive or linguistic factors.

Predictors of Reading Comprehension

Research Question 4 asked how accurately reading comprehension can be predicted (a) from IQ, (b) from a set of linguistic comprehension measures (i.e., vocabulary and listening comprehension) over and above IQ, (c) from decoding rate over and above IQ and linguistic comprehension measures, and (d) from the combination of IQ, linguistic variables, and decoding rate. To answer this question, a sequential multiple regression was performed to assess the predictive value of each of the specified measures individually and in combination.

Statistically significant correlations were obtained between reading comprehension and all the predictor variables. The correlations were moderately high: Those between reading comprehension and the linguistic measures exceeded .50 and that between reading comprehension and decoding approached .40. Together, the predictors accounted for more than 40% of the variance in reading comprehension, with IQ explaining most of the variance, 29%, and linguistic measures (i.e., listening comprehension and vocabulary) explaining an additional 13%. Among the linguistic measures, the contribution of vocabulary was not statistically significant and its relationship with reading comprehension appeared to be fully mediated by the relationship between reading comprehension and listening comprehension. The contribution of decoding rate failed to reach statistical significance.

Previous studies in English (e.g., Joshi & Aaron, 2000; Swanson et al., 2003; Tunmer & Hoover, 1993) have reported a substantial association between

word decoding and reading comprehension, with correlations often in the .60-.70 range and exceeding those between reading comprehension and verbal ability, IQ, or memory. The results of the present study, however, do not appear to support such a large association. Does this mean that decoding ability is unimportant in determining reading comprehension in the Russian language? I believe that such a conclusion would be premature.

First, the correlation between decoding rate and reading comprehension obtained in the present study was statistically significant and moderate in magnitude, at .37. Similar correlations have sometimes been observed in other languages including English (Hoover & Gough, 1990) and Turkish (Oney & Goldman, 1984), where decoding was found to be a reliable predictor of reading comprehension. Second, it seems reasonable to assume that a complete absence of decoding ability will result in no reading comprehension regardless of the individual's level of linguistic comprehension or IQ, so in this sense, decoding *is* a necessary condition for reading comprehension. Third, the discriminant function analyses performed in this study to determine cognitive abilities that distinguished reliably between good and poor readers clearly indicate that decoding *is* relevant to reading comprehension. So why didn't decoding ability contribute independent variance to reading comprehension?

I believe that the most plausible explanation is methodological. In determining the appropriate sample size needed for this study, I estimated, based on prior research, the degree to which the criterion variable, reading comprehension,

was related to the predictor variables in the population (i.e., effect size). In the absence of reliable research on reading comprehension in the Russian language, I used studies conducted in languages with less inconsistent orthographies than English (e.g., Dutch, French, Turkish) to estimate effect size. However, such an estimation was bound to be approximate and not definitive. I therefore believe that the size of the sample used in this study was not sufficiently large to bring out the contribution of decoding, although in the regression analysis, it approached statistical significance (p = .07).

If this interpretation is correct, it provides additional evidence for the claim (e.g., de Jong & van der Leij, 2002; Megherbi et al., 2006) that the relative weight of abilities that contribute to reading comprehension may depend on the characteristics of the particular orthography. Russian orthography is largely transparent for reading and native Russian children appear to master the challenges of word decoding very fast. Most of the children in the sample were proficient decoders, as revealed by the Rasch analysis of the decoding test, and more than half could read real words at or above the prescribed level of 30 wpm. Thus, it appears that the children in the sample had already approached—and some had undoubtedly crossed—the critical threshold of decoding above which this process becomes largely automatic and individual differences in decoding cease to be crucial to reading comprehension. At this stage, the contribution of individual differences in decoding to reading comprehension may be so small that it cannot be detected after the differences in other cognitive abilities are accounted for.

In contrast to decoding skills, linguistic comprehension was a powerful predictor of reading comprehension. This result is consistent with previous research in English and other alphabetic languages (see the Literature Review). What is interesting, however, is that the effect of listening comprehension turned out to be much greater than the effect of vocabulary. In fact, as the regression analysis revealed, the former fully accounted for the latter. A possible explanation for this result can be found in the nature of the vocabulary test used in this study. The test, which is part of the *Wechsler Intelligence Scale for Children*, Russian Version (Philimonenko & Timofeev, 2001), is a test of expressive vocabulary measuring not only the ability to comprehend word meanings but also the ability to provide definitions. The listening comprehension test, on the other hand, was a purer measure of linguistic comprehension ability as it did not rely heavily on the clarity of expression. In fact, to the extent that decoding ability no longer exerts an additional influence on reading comprehension, both listening and reading comprehension can be conceived as tapping the same ability, the ability to comprehend language, oral and written. The listening test also served as an indirect measure of receptive vocabulary, which has been shown to be a strong predictor of reading comprehension (Goff et al., 2005). The presence of a receptive vocabulary component in the listening test may also explain why it was highly correlated with the vocabulary test.

The results of the multiple regression analysis also indicate the importance of IQ to reading comprehension. In the present study, IQ had a substantial

correlation with reading comprehension, exceeding that between reading comprehension and vocabulary and reading comprehension and decoding rate. Furthermore, IQ explained a large portion of variance in reading comprehension when entered first in the regression equation. This result is consistent with previous findings in English showing IQ to be a moderate predictor of reading comprehension (e.g., Naglieri, 2001; Stanovich et al., 1984). An important implication of this finding is that IQ should be incorporated into models of reading comprehension—including those based on the Simple View—to account for individual differences in nonverbal intelligence.

What is interesting is that IQ, linguistic comprehension, and decoding rate could not account for all the variance in reading comprehension. This means that other abilities not involved in this study are likely to influence the development of reading comprehension in Russian children. One such candidate ability is processing speed. Indeed, the inclusion of processing speed measures has been shown to improve the predictive power of the Simple View formula (Joshi & Aaron, 2000). Other possibly relevant abilities include working memory (Goff et al., 2005; Swanson & Berninger, 1995) and the use of context, which were not measured in this study.

Cognitive and Linguistic Abilities of Good and Poor Readers

Research Question 5 asked about the differences between good and poor readers in the population on linear combinations of IQ, phonological awareness,

listening comprehension, vocabulary, and reading comprehension and whether the children in the sample could be correctly classified as good and poor readers based on their scores on the five predictor variables. Two sets of direct discriminant function analyses were performed, one with the whole sample split into two groups and the other with the good and poor readers separated by a buffer zone. In both sets of analyses the groups differed reliably on all five predictors, with reading comprehension, phonological awareness, and vocabulary being the strongest differentiating variables. The explanatory power of the discriminant function was much greater in the more extreme groups where it explained more than twice as much variance as in the less extreme groups (42% vs. 18%).

Further, in both analyses, there were strong correlations between vocabulary and IQ, vocabulary and phonological awareness, and vocabulary and reading comprehension, indicating that children with strengths in vocabulary were likely to have higher nonverbal intelligence, better reading comprehension ability, and stronger phonological awareness skills. The magnitude of the correlations was greater in the more extreme groups. The classification accuracy for the original sample as well as for a cross-validation run was moderate (kappa coefficient = .30) for the less extreme groups and high (kappa coefficient = .63) for the more extreme groups.

The analyses show a clear pattern in the profile of strengths and weaknesses of young readers of Russian: The better a child's word decoding skills, the more likely that child is to comprehend what they are reading, demonstrate efficient

access to phonology, and have a larger vocabulary. Moreover, children with stronger decoding skills are likely to have higher IQs and better listening comprehension ability, although the differences between good and poor readers on these variables are smaller than on the other three variables. Obviously, because of the correlational nature of the present study it is impossible to say with confidence which cognitive abilities are causally related to reading. Nevertheless, several tentative conclusions can be drawn from the patterns of association between decoding and cognitive skills uncovered in this study.

First, the consistent differences between good and poor readers on reading comprehension observed in the present study highlight the importance of decoding to individual differences in reading comprehension. Children with deficits in decoding ability are more likely to have problems with reading comprehension than children who are proficient decoders. This finding underscores the importance of bottom-up skills in reading, indicating that these skills may very well drive the development of the reading system. It also strengthens the methodological explanation provided earlier in this section for the failure of the multiple regression analyses to reveal a unique contribution of decoding to reading comprehension.

Second, the reliable differences between good and poor readers on vocabulary observed in the present study indicate the special role vocabulary plays in both reading comprehension and word decoding. Such a role would be consistent with recent findings in English-language research showing vocabulary to be not only a unique predictor of both reading comprehension and word recognition

(Nation & Snowling, 2004) but also a path to reading comprehension for some children (Berninger et al., 2006). It is also compatible with Ehri's (2002) view that fluent word recognition depends on the reader's familiarity not only with the word's pronunciation but also with its meaning.

Third, the differences between good and poor readers on phonological awareness are consistent with the earlier result obtained in this study showing that phonological awareness plays a role in determining decoding fluency in young Russian readers. This result further contributes to the growing literature on the importance of phonological awareness to the development of decoding ability in all alphabetic languages. Furthermore, phonological awareness in this study showed high correlations with vocabulary, indicating that it probably plays a role in the development of verbal skills in young readers. It can be speculated, for example, that an underdeveloped phonological ability would impede the development of an effective system for mapping speech sounds onto units of meaning and interfere with an efficient storing of new phoneme sequences in memory, thereby constraining vocabulary growth. In her review of research on German-speaking children with dyslexia, Landerl (2003) came to a similar conclusion.

Finally, the differences between the two groups on IQ, together with the strong intercorrelations observed among IQ, reading comprehension, and vocabulary appear to suggest that nonverbal ability plays an important role in both reading comprehension and vocabulary development. This result is consistent with previous research showing the importance of nonverbal ability to reading

comprehension (e.g., Tiu et al., 2003) as well as a strong association between this ability and vocabulary skills (e.g., Goff et al., 2005). It also strengthens the findings of the multiple regression analyses on the role of IQ in reading comprehension.

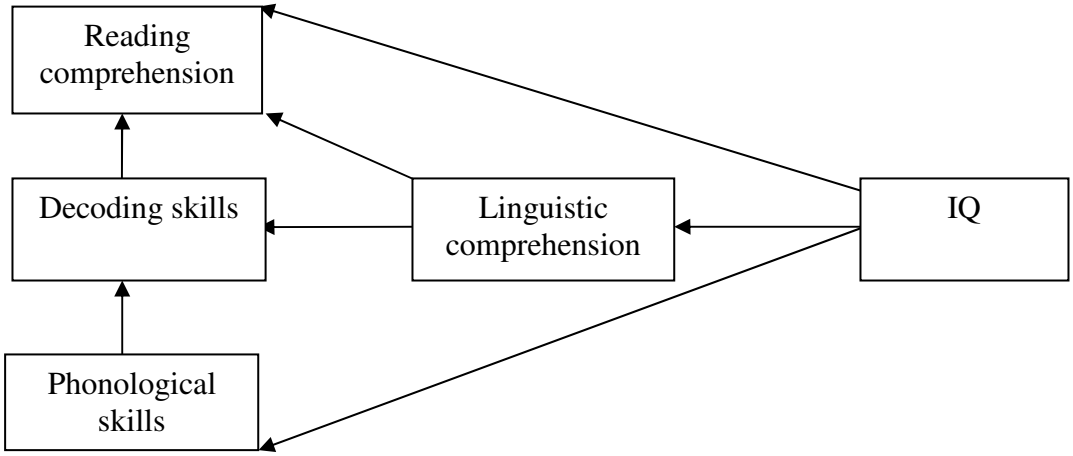


Figure 9. A conceptual model of reading comprehension in Russian.

Taken together, the results of the multiple regression and discriminant function analyses suggest a tentative model of reading comprehension in young Russian readers. This model is shown graphically in Figure 9. As shown in the figure, phonological skills (e.g., phonological awareness and RAN) drive the development of efficient decoding skills, which in turn provide the foundation for reading comprehension. However, decoding alone is not sufficient for reading comprehension as meaning comprehension is also required. The importance of meaning comprehension to reading is captured in the connection of linguistic

comprehension (i.e., listening comprehension and vocabulary) not only to reading comprehension but also to decoding. Finally, the model specifies a role for IQ in the development of reading comprehension, linguistic skills, and phonological skills and provides indirect support for Turner's (1997) observation that many reading problems may be a result of low intelligence.

This model also supports the Simple View of reading, in which reading comprehension is a function of both decoding ability and linguistic comprehension and shows that a deficiency in decoding will lead to deficits in reading comprehension. However, the results of the present study clearly indicate that the relationship between decoding and linguistic comprehension does not exhaust all of the variance in reading comprehension ability. Furthermore, the unidirectional influences of cognitive abilities on reading specified in this model are clearly speculative and are equally likely to be reciprocal or even in the opposite direction, at least for some abilities. More research is needed before a more definitive conclusion can be drawn about the specific role of cognitive abilities in reading in Russian.

Cognitive and Linguistic Profile of the Poorest Reader

Research Question 6 asked about the cognitive, linguistic, and reading profile of the poorest readers in the sample. Only one child, a girl (NY), matched the specified definition of extremely poor reading, scoring more than 2 *SD* below the mean on both decoding accuracy and decoding rate. On the word reading test,

NY was unimpaired on accuracy but showed marked deficits in speed. On the nonword reading test, however, she was severely impaired on both speed and accuracy, taking almost 13 minutes to read 48 items and scoring 7 *SD* below the mean on decoding speed. Her performance on phonological awareness, short-term memory, IQ, and vocabulary was unimpaired, whereas her performance on the reading and listening comprehension tests was close to that of the poor readers. The cognitive variable that most clearly separated NY from both the good and poor readers was RAN rate.

Several conclusions can be drawn from NY's reading performance. First, it is clearly possible for children with weak phonological recoding skills to master word reading in the Russian language given enough exposure to printed words. Moreover, such children can achieve high accuracy in word reading even though they may be unable to map precisely between the letters on a page and the sounds they represent. Despite her severe impairment in phonological recoding as indicated by her performance on nonword decoding accuracy, NY was able to demonstrate average performance on word reading accuracy.

Second, consistent with previous research in transparent orthographies (Porpodas, 1999; Wimmer, 1993), NY's deficit in word reading manifested itself primarily in effortful and slow decoding. Moreover, a comparison of NY's performance on morphologically simple and morphologically complex words suggests a possible trade-off between accuracy and speed in her performance: On the list of simple words, she scored 100% on accuracy but was 4.4 *SD* below the

mean; in contrast, on the list of more complex words, she scored 70% on accuracy but was only 1.5 *SD* below the mean. This result suggests that accurate word reading in transparent orthographies is achieved by children with literacy difficulties at the cost of slow and laborious processing of print.

What is puzzling, however, is NY's apparently complete inability to read *any* of the nonwords, including the simplest CVC items. Yet many previous studies (e.g., Porpodas, 1999; Wimmer, 1993; 1996) conducted in languages with transparent orthographies have shown that children with dyslexia *can* read at least some of the nonwords accurately, although they almost always fall behind average readers on this task. How, then, could NY's utter failure with nonword reading be explained? More important, how could this failure be reconciled with NY's seemingly unaffected word reading accuracy? I believe that a plausible explanation lies in the notion of individual differences in dyslexia (Rack et al., 1992; Snowling, 2000). According to this view, individual differences between dyslexics depend on both the severity of their core phonological deficit and the interaction of this cognitive deficit with individual strengths and weaknesses. It is therefore possible that NY's reading performance reflects her individual, idiosyncratic style and use of strategies resulting from her own pattern of cognitive strengths and weaknesses.

For example, it can be hypothesized that NY's phonological deficit was so severe that it prevented her —at least in the time that had passed since she began learning to read—from forming fine-grained phonological representations needed for the precise mapping between sound and print. Her apparent deficit in the

automaticity of lexical access as revealed by her poor performance on the RAN test (see Table 26) could have further exacerbated the problem, making verbal retrieval difficult. These cognitive deficits—perhaps coupled with insufficient print exposure—would then explain her inability to read nonwords. On the other hand, NY's above average IQ and intact STM and vocabulary skills could have made it possible for her to acquire, through the memorization of common letter patterns and their pronunciations, a small sight vocabulary of high-frequency words. She would then use these word-specific associations to visually recognize familiar words. In this sense, NY's representations of words would be coarse, underspecified, and lexical-visual rather than sublexical-phonological, and her reading strategies would be based on a visual approach.

To put this hypothesis to the test, I compared NY's reading errors on the two types of nonword used in this study—those that were visually similar to real words and those that were visually dissimilar from real words. If this hypothesis were correct, then visual similarity should make a difference to NY's reading performance, resulting in her errors on the visually similar nonwords being more systematic. Specifically, I expected that on these nonwords, her reading errors would be substitutions of the target items with similar-looking real words; on the visually dissimilar nonwords, however, her reading errors would be much more random. Moreover, she would need more time to find a plausible real-word alternative for a visually dissimilar item than she would for a visually similar one, which would result in larger naming latencies for the dissimilar items.

A qualitative analysis of NY's reading errors on the two types of nonword stimuli confirmed these expectations. On the visually similar nonwords (n = 24), she substituted 11 target items with the original words on which the nonwords were based and the remaining 13 with very similar-looking alternatives (see Table 27). In contrast, her errors on the 24 visually dissimilar nonwords were much less systematic and in most cases, NY produced a much shorter word that resembled the target item only in the first few letters. These errors suggest that she probably tried to use a partial visual strategy to read the nonwords. A comparison of NY's timed performance on both types of nonword supports this conclusion: Although she was equally inaccurate on both tests, NY could generate a pronunciation for a visually similar nonword much faster than she could for a visually dissimilar one (see Table 25). Arguably this is because finding an acceptable real-word alternative for a visually similar nonword takes less time than finding a suitable real-word alternative for a dissimilar nonword, particularly one with an uncommon letter pattern.

For decoding accuracy, however, this visual strategy was equally ineffective for similar and dissimilar nonwords because such words are not present in one's mental lexicon and hence require precise mapping between print and sound for accurate decoding. The use of the visual strategy may also explain why NY was so successful on word reading accuracy: Because all the words on this test came from age-appropriate textbooks, they were likely to be present in young readers' mental lexicon and could be recognized accurately even from coarse phonemic cues.

The striking dissociation between NY's word and nonword reading performance and the dramatic difference observed in her reading of the visually similar versus visually dissimilar nonwords suggests a certain degree of independence between the processes of phonological recoding and sight word recognition in Russian, at least for some children. This conclusion is supported by evidence from English-language studies showing that many dyslexics are able to acquire word recognition skills well beyond the level that would be expected based on their phonological skills (e.g., Snowling, Stackhouse, & Rack, 1986). It should be emphasized, however, that the conclusions regarding NY's reading performance are speculative and should be interpreted with caution because NY's visual skills were not tested directly in this study.

Table 27. *NY's Reading Errors*

Visually similar nonwords	NY's reading	Visually dissimilar nonwords	NY's reading
юскра	искра (spark)	контрапонист	контора (office)
встроча	строчка (line)	блонтерстапинг	балон (tank)
конфата	конфета (candy)	безарё	без (without)
дровосяк	дровосек woodcutter)	пенерифул	пенек (stump)

It is also noteworthy that in absolute terms, it is not uncommon for non-English dyslexics to show poor performance on nonword reading even after several years of schooling. For example, Wimmer (1996) found that after four full years of reading instruction, German-speaking dyslexics could read only about 80% of the

relatively simple open-syllable nonwords such as *toki* and *kerata* correctly. NY was tested after just one year of reading instruction and most of the items on the nonword reading test were orthographically complex. It is possible that this time was simply insufficient for her to establish more precise phonological representations needed for the processing of print at the phonemic level.

NY's cognitive and linguistic profile was similar to that of the average readers in the sample with two exceptions: She was impaired on reading and listening comprehension relative to the normal reader group and severely impaired on rapid naming relative to both the good and poor readers. In fact, her performance on RAN was more than 2 *SD* below the mean of the good readers and more than 1 *SD* below the mean of the poor readers. This result is in agreement with previous findings showing that in transparent orthographies, RAN is one of the strongest predictors of literacy problems (see the Literature Review).

What is particularly remarkable is that NY demonstrated average performance on all phonological awareness tasks. And yet her severe impairment on nonword reading clearly indicates that her phonological skills could not have been intact. Evidently the orthographic transparency of the Russian language must create favorable conditions for the development of phonological awareness so that even children with phonological deficits can acquire at least some phonological skills during the first year of reading instruction. This finding is consistent with previous research showing the difficulty of detecting phonological impairments in poor readers and readers with dyslexia in transparent orthographies. Nikolopoulos

and his colleagues (2003), for example, have argued that in such orthographies, difficulties with phonology can be uncovered only with the more cognitively demanding tasks such as those requiring phoneme transposition. This result highlights the need voiced by many researchers (e.g., Goulandris, 2003a) to consider reading disability in terms of the specific language of instruction because its behavioral manifestations can only be identified with reference to the core characteristics of the spoken and written language in use.

In summary, it appears that in Russian, a relatively transparent orthography, word recognition can be achieved even with weak phonological skills because coarse phonological representations suffice for accurate word recognition although the weakness in phonology results in considerable delays. In contrast, nonword decoding, which relies heavily on phonological processing and requires fine-grained phonological representations becomes, in the absence of a well-developed phonological system, an arduous and error-prone task. The findings also indicate that measures of word reading accuracy would be of limited value in assessing young children's reading performance. The crucial indicators of children's difficulties in processing Russian orthography appear to be the time they take to read both words and nonwords and the accuracy they demonstrate in decoding nonwords. Furthermore, simple phonological tasks may be insufficient to uncover problems with phonology in the Russian language, although poor performance on RAN may indicate underlying cognitive deficits.

Concluding Speculations: Implications of the Rasch Analysis

The Rasch analysis conducted in the present study makes it possible to assess the psychometric properties of the two commercial testing instruments used, the Vocabulary Subtest of the *Wechsler Intelligence Scale for Children*, Russian Version (WISC, Philimonenko & Timofeev, 2001) and the *Kaufman Brief Intelligence Test,* 2nd ed. (KBIT-2, Kaufman & Kaufman, 2004). The usefulness of these tests is discussed below.

The Vocabulary Subtest of the WISC is a popular measure of verbal ability in Russia, where it is used for both research and practical purposes such as to diagnose developmental delays (Kornev, 2006). However, the results of the present study suggest that the test in its present form may be inappropriate for children in early elementary grades. A large discrepancy was revealed between the level of difficulty of the test items and the level of ability of the children in the sample, with the mean item difficulty being more than 2 *SD* above the mean person ability. The failure of many of the children in the present study to define clearly even the simplest words on the test such as *knife, bicycle*, and *pillow* is particularly worrisome as it suggests that the assumptions on which this test is based do not reflect the level of ability and knowledge that young Russian children can be realistically expected to demonstrate. Further, although the fit statistics and the PCA of standardized residuals showed that the test tapped a unidimensional construct, the fit statistics revealed a large number of overfitting items (more than 30% of the test), which indicates that at least a third of the items on this test

provide a limited amount of useful information about the children's vocabulary knowledge.

In contrast, the results for the KBIT-2 indicate that this test can be a useful tool for obtaining a quick estimate of a child's intelligence, for example, for screening purposes to identify children who may benefit from remedial programs. The test appears to be well targeted for children in the early elementary grades, and it features high reliability and very high internal validity as revealed by the PCA of standardized residuals. However, as shown in the Rasch item-person map (see Figure 7), the test may benefit from some modification for use with young school children. For example, it could be shortened by eliminating the first section and reducing the length of the last section, and the order of the items could be re-arranged to reflect the progression of difficulty revealed by the Rasch analysis.

An important finding revealed in this study by the Rasch analysis is that word and nonword decoding represent essentially the same construct. Previous studies have often treated word and nonword reading as different measures—the former as a measure of word recognition and the latter as a measure of phonological recoding and have even used nonword reading to predict word reading. The justification for such treatment comes primarily from dual-route theories of reading (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), which assume the existence of two distinct paths by which readers can arrive at a word's pronunciation: a phonological path, which relies on letter decoding, and a direct visual path, which relies on visual-

orthographic processing in lexical access. According to these theories, real words can be read using both a phonological and a visual strategy whereas nonwords, which presumably are not contained in lexical memory, can only be read by phonological recoding. Thus, nonword reading becomes a test of the integrity of the phonological system.

However, the present study does not support the existence of two separate paths to meaning, at least for normally developing children learning to read a transparent orthography such as Russian. As was revealed by the Rasch analysis, both the word and the nonword reading tests were closely related and both appeared to tap the same underlying ability. This means that at least for normally developing readers of Russian, word reading, just as nonword reading, is achieved primarily through phonological recoding, which readers in the present study used with various degrees of success.

This finding provides further support for the psycholinguistic grain-size theory (Ziegler & Goswami, 2006), which stipulates differential development of the reading system in different orthographies. It also cautions against the use of nonword reading to predict word reading in normally achieving children learning to read a transparent orthography, because these abilities appear to be largely equivalent and their unique predictive variance would hardly be informative. Investigations into the relationships between cognitive tasks and predictors of reading ability must take into account the fact that both word and nonword reading tests tap the same latent construct, at least in regular orthographies.

CHAPTER 7

CONCLUSION

Summary of the Findings

The present study investigated the concurrent predictors of decoding and reading comprehension in young Russian readers in order to identify the strongest independent predictors and to determine to what extent predictors of reading in Russian are similar to those in other languages, most notably in English. The structure of common phonological awareness tasks was also examined to determine if they tap the same underlying ability in Russian. Furthermore, the cognitive and linguistic abilities of good and poor Russian readers were explored to determine the largest sources of difference. The particular strength of the present study is in the use of truly objective, highly reliable, interval-scale measures of ability used to represent the main constructs. These measures allow me to have confidence in the results, which is further bolstered by the convergence of these results with those of other multivariate studies.

The present study supports the findings from other alphabetic languages, particularly those with transparent orthographies, showing that the progress children make in learning to read is related to individual differences in both phonological awareness and RAN skills and extends these findings to the Russian language. The finding that phonological awareness was also a reliable differentiator between good and poor readers further supports the important role of phonology in

fluent reading development in transparent orthographies. The similarities between the cognitive predictors of reading observed in this study of Russian children and those observed in other transparent languages suggest that models of literacy development, and hence theories of literacy deficits, may, to some extent, generalize across alphabetic languages.

Yet, the present results also indicate that the role of cognitive abilities that affect reading development and their relative importance in predicting reading achievement may depend, at least in part, on the features of the orthography a child is learning. Specifically, the role of phonological skills in predicting Russian word reading, although important, appears to be much smaller than that typically observed in English and less substantial than the role of rapid naming. Further, verbal short-term memory as measured in this study was found to be unrelated to reading ability. These results are attributed to the effect of the Russian orthography, which appears to enable direct print-to-sound conversion thereby permitting, in the presence of sound-letter knowledge, the sequential assembly of word pronunciations. As such, these results provide additional support for the role of language-specific factors in the development of the reading system and highlight the importance of incorporating mechanisms capable of accounting for language-specific differences into existing reading models.

With regard to reading comprehension, partial support was found for a positive connection between the efficiency of word decoding and the accuracy of reading comprehension. This support came in the form of a significant correlation

observed between decoding and reading comprehension as well as a reliable difference that emerged between good and poor decoders on reading comprehension. However, in the multiple regression analyses, decoding failed to account for additional variance in reading comprehension after IQ and linguistic comprehension were included in the regression equation. This result is interpreted as further evidence for the role of the writing system in determining the relative weight of abilities that are important to reading in a given language. The straightforward nature of the Russian script must have made it possible for many children in the sample to acquire, just after one year of reading instruction, a level of decoding sufficient for accurate word recognition and as a result, decoding stopped exerting an additional influence on reading comprehension.

The results of the present study also highlight the importance of linguistic abilities to both word reading and reading comprehension. Linguistic comprehension was a reliable predictor of reading comprehension even when the differences between the children in IQ were statistically controlled. On the other hand, children who were better readers tended to have larger vocabularies and better listening comprehension skills. Further, nonverbal ability was found to be a significant predictor of reading comprehension, accounting for a substantial portion of variance. These results indicate that the operation of both word reading and reading comprehension may, to some extent, be shaped by general language competence as well as general intelligence, at least in children who have achieved a certain level of decoding skill. They also suggest the need to incorporate measures

of nonverbal intelligence—and possibly vocabulary—into reading models based on the Simple View of reading.

Another important finding concerns the nature of phonological awareness ability. The results of the Rasch analysis indicate that the phonological tasks of rhyme and phoneme detection, phoneme segmentation, and phoneme deletion used in the present study tap the same underlying ability in the Russian language. These findings add to a growing literature in English and other languages showing that phonological awareness is best conceptualized as a single ability that can be measured by tasks of varying linguistic complexity and cognitive demands. For practical purposes, this means that phonological awareness can be accurately assessed even in children who have not yet begun to read.

Further, the present study provides a glimpse into the patterns of difficulty dyslexic readers may exhibit in the Russian language. Just like in many other transparent languages, in Russian, too, dyslexia appears to present primarily as a speed deficit in both word and nonword reading and as an accuracy deficit in nonword reading. On the other hand, the relative transparency of the Russian orthography appears to permit even readers with phonological deficits to achieve a high level of accuracy in word reading. Further, the intact performance of the dyslexic child on the phonological battery observed in the present study underscores the difficulty of identifying phonological deficits in readers of a transparent orthography. These results raise doubts about the appropriateness of generalizing studies of dyslexia predominantly conducted in English to other

orthographies. They also highlight the need for creating language-appropriate tools for assessing literacy problems.

From a practical perspective, these results suggest that in the Russian language, reading speed may be the more appropriate measure of reading achievement than reading accuracy, although nonword reading accuracy may also indicate literacy problems. Further, RAN tasks are probably the most sensitive measure to use in screening programs to identify children with literacy problems. Phonological awareness tasks may also be helpful in differentiating good from poor decoders, although they may be of limited value with at least some dyslexic individuals.

Finally, the findings of the present study provide indirect support for the psycholinguistic grain-size theory of reading (Ziegler & Goswami, 2005), which stipulates differential development of reading strategies in response to the nature of the orthography a child has to master, particularly the degree of its transparency. This support comes primarily from the Rasch analysis of word and nonword reading tests, showing that both tests tap the same underlying ability in normal readers of Russian. Contrary to the findings of English-language research, which has documented several separate routes to meaning in word reading including reading by analogy and reading by whole-word strategies, in Russian, lexical access in reading appears to rely almost exclusively on phonological recoding, at least in young average-ability readers. This finding supports the conclusion that the

nature of word reading may be qualitatively different in consistent versus inconsistent orthographies, at least at an early stage of reading development.

Limitations and Suggestions for Future Research

Several characteristics of the present study limit the generalizability of the obtained results. First, the correlational nature of the present analyses prohibits the attribution of causality to the results and precludes any causal inference about the relationship between the cognitive and linguistic abilities measured in this study and reading achievement. Although phonological awareness and RAN were found to be reliable predictors of decoding whereas linguistic abilities and IQ were found to be reliable predictors of reading comprehension, the exact role of these abilities in determining the reading outcome remains speculative. Training studies can help disambiguate the nature of some of the relations observed in the present study, particularly between phonological awareness, rapid naming, and decoding and between linguistic and reading comprehension.

Another important limitation of this study is its concurrent nature. Although the present study provides a first glimpse into the process of reading acquisition in the Russian language, reading was investigated here only at one point in time. Recent longitudinal research in English and other languages, however, has revealed the time-limited effects of many cognitive abilities on reading achievement and has demonstrated that the relationship between reading and various cognitive and linguistic abilities related to it changes over time as a function of age, teaching, and

the amount of print exposure. Under these circumstances, longitudinal studies of reading become particularly important as they can clarify how the relationship between decoding and reading comprehension on the one hand and nonverbal ability, phonological awareness, rapid naming, verbal memory, and linguistic ability on the other changes in the course of development.

Third, the obvious limitation of this study is the relatively small sample size. Although the sample size was considered adequate based on the preliminary estimation of effect size for the predictor variables, this estimation was based on studies in languages other than Russian and as such, was bound to be approximate. The insufficiently large sample size may account for the failure to obtain stronger correlations between cognitive and reading measures in a number of analyses. Clearly the findings require replication with a larger sample. Nevertheless, the high reliability and internal validity of the truly continuous measures employed in the present study as well as the convergence of many of the present findings with those that have been reported previously for other transparent languages allows me to have high confidence in these results.

Further, this study was conducted in Moscow, which is not only the biggest city in Russia but also one that has traditionally been viewed as the most desirable for residence, with better schools, better health-care, and better social infrastructure. It is therefore difficult to generalize results obtained with Moscow children to all Russian children. Clearly findings obtained in less developed and less affluent geographical locations would be different, at least to some extent. An

important direction for future work is to investigate reading development in children living in non-Moscow regions. This is particularly important if we are to create reliable standardized assessment instruments to gauge reading progress and reading development of children across Russia. Nevertheless, the present results provide a first glimpse into the pattern of normal reading development in the Russian language from the cognitive perspective and as such may generalize to other average-ability children living in large Russian cities.

Another potential limitation arises from the test administration procedures used in this study. Because the tests took an average of 90 minutes per child, an obvious concern is the presence of maturation effects including fatigue, boredom, and loss of motivation. Furthermore, since some of the tests were administered by an assistant, it can be argued that the differences in style and personality may have introduced inadvertent bias in the outcome. However, while acknowledging that it is hardly possible to eliminate the risk of bias completely under such testing conditions, I believe that any potential bias is unlikely to have been substantial for the following reasons. First, the tests were ordered in a way that allowed alternating cognitively demanding tasks with easier ones in order to maintain the child's interest and motivation. Second, the children were constantly monitored for signs of fatigue or boredom and were encouraged to take a break if they felt they needed one. Most of the children appeared at ease during testing and they seemed happy to participate. Finally, to minimize the chances of bias due to the differences in style between the examiners, all tests that relied heavily on auditory processing (e.g.,

phonological, RAN, STM, and listening comprehension tests) were administered by the same examiner and under the same voice conditions to all the children.

Finally, the present study has focused on cognitive and linguistic factors of reading achievement. However, it is clear that other, non-cognitive factors may affect the reading outcome as well as the process of reading development. These factors may include, among other things, the teaching method, the amount of print exposure a child has had, and family encouragement. Another important factor that was not investigated in the present study is the use of context in reading and its effect on both word decoding and reading comprehension. In the present study, reading was measured using word lists to obtain a purer measure of decoding ability. However, in real-life situations, reading tasks usually involve the use of context, which can significantly facilitate both word recognition and meaning comprehension, particularly for poor readers. Future research should focus on these unexplored areas of reading and reading comprehension to tease apart the contribution of these factors to the concurrent and longitudinal development of reading skills among Russian children.

Conclusion

The present study is a first attempt to investigate systematically the role of a specific set of cognitive and linguistic abilities in the concurrent prediction of word decoding and reading comprehension in the Russian language. As such, it extends many of the findings of previous research to the Russian language. The analyses

revealed similarities between the main psychological operations involved in Russian reading and those that have been described in other languages, including English. These similarities include the importance of phonological and RAN skills to word decoding, the role of general intelligence and linguistic comprehension in determining reading comprehension, and the manifestations of reading disability. These results provide further support for the conclusion that reading in all alphabetic languages may involve similar cognitive and metalinguistic mechanisms. However, the present study also uncovered a number of important characteristics that make Russian reading distinct from reading in other languages. These characteristics concern the involvement of verbal short-term memory in decoding, the role of decoding in reading comprehension, and the pattern of cognitive and reading skills exhibited by poor readers. Taken together, the findings of the present study support the view that reading in all alphabetic languages is characterized by both universal and language-specific features. Furthermore, they provide the foundation for the construction of language-appropriate models of reading development and reading disability in the Russian language.

This study has made several important contributions to the scientific study of reading. First, it has clarified the relative contribution of a number of cognitive and linguistic abilities to reading development in the Russian language, thereby making it possible to place Russian reading within the framework of existing theories of literacy development and cross-linguistic reading research. Second, it has clarified the measurement properties and psychometric usefulness of a battery

of tests designed to measure cognitive abilities and reading achievement in young Russian children and has prepared the ground for the development of culture-appropriate, theoretically justified, and scientifically objective reading and cognitive assessment tools. Third, it has confirmed the unidimensional structure of phonological awareness ability, thereby contributing to the resolution of a long-standing controversy over which phonological units are more important to reading development, and has shown that this ability can be assessed by a variety of linguistically and cognitively diverse tasks. Finally, the present study has clarified the main sources of difference between good and poor Russian readers, described a pattern of difficulty Russian dyslexic readers may exhibit in the course of reading acquisition, and indicated which tests might be most effective in the detection of literacy difficulties in the Russian language.

REFERENCES CITED

Aaron, P. G., & Joshi, R. (1992). *Reading problems: Consultation and remediation*. New York: Guilford.

Abu-Rabia, S. (2004). Dyslexia in Arabic. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross language comparison and practice guide* (pp. 31-38). Chichester, UK: John Wiley.

Abu-Rabia, S., Share, D., & Mansour, M. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing: An Interdisciplinary Journal, 16,* 423-442.

Adams, M. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Akhutina, T. (2002, June). *A. R. Luria approach to prevention and remediation of writing difficulties.* Paper presented at the ISCRAT Congress symposium, Amsterdam, the Netherlands.

Akhutina, T. (2003). L S. Vygotsky and A. R. Luria: Foundations of neuropsychology. *Journal of Russian and East European Psychology, 41*(3/4), 159–190.

Akimova, M., & Kozlova, V. (2006). *Diagnostika umstvennogo razvitiya detei* [Diagnosis of intellectual development in children]. Moscow: Piter.

Alekseenko, M., Bantova, G., Beltukova, G., Velichkina, S., Volkova, M., Golovanova, et al. (2003). *Shkola Rossii* [School of Russia]. Moscow: Prosveshenie.

Algina, J., Moudler, B., & Moser, B. (2002). Sample size requirements for accurate estimation of squared semi-partial correlation coefficients. *Multivariate Behavioral Research, 37*(1), 37-57.

Anthony, J., & Francis, D. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*(5). 255-259.

Anthony, J., & Lonigan, C. (2004). The nature of phonological sensitivity: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology, 96*, 43-55.

Anthony, J., Lonigan, C., Burgess, S., Driscoll, K., Philips, B., & Cantor, B. (2002). Structure of preschool phonological sensitivity: Overlapping

sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology, 82*, 65-92.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*, 621-635.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (1992). Is working memory working? The fifteenth Barlett lecture. *The Quarterly Journal of Experimental Psychology Section A, 44*(1), 1-31.

Ball, E., & Blachman, B. (1991). Does phoneme training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly, 26(*1), 49-66.

Bentin, S., & Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: The case of Arabic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(2), 309-323.

Berninger, V., Abbott, A., Vermeulen, K., & Fulton, C. (2006). Paths to reading comprehension in at-risk second-grade readers. *Journal of Learning Disabilities, 39*(4), 334-351.

Besner, D., & Smith, M. (1992). Basic processes in reading: Is the orthographic depth hypothesis sinking? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 45-66). Amsterdam: North-Holland.

Bialystok, E., McBride-Chang, C., & Gigi, L. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology, 97*(4), 580-590.

Bode, R. (2004). Partial credit model and pivot anchoring. In E. Smith & R. Smith (Eds.). *Introduction to Rasch measurement: Theory, models, and applications* (pp. 279-295). Maple Grove, MN: JAM Press.

Boldyreva, T., & Inshakova, O. (2004). Dyslexia in Russia. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross-language comparison and practice guide. Part 2: Countries* (pp. 197-200). Chichester, UK: Wiley.

Bond, T., & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.

Bowers, P., Steffy, R., & Tate, E. (1988). Comparison of the effects of IQ control methods on memory and naming speed predictors of reading disability. *Reading Research Quarterly, 23*(3), 304-319.

Bowey, J. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63*, 44-78.

Bradly, L., & Bryant, P. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 301*, 419-421.

Bruck, M. (1988). The word recognition and spelling of dyslexic children. *Reading Research Quarterly, 23*(1), 51-68.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*(5), 874-886.

Bus, A., & Ijzendoorn, M. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*(3), 403-414.

Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly Journal of Experimental Psychology Section A, 37*(3), 435-475.

Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology, 55*, 1-30.

Cardoso-Martins, C., & Pennington, B. (2004). The relationship between phoneme awareness and rapid serial naming skills and reading acquisition: The role of developmental period and reading ability. *Scientific Studies of Reading, 8*(1), 27-52.

Catts, H., Hogan, T., & Fey, M. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36*(2), 151-164.

Chernyavskaya, (2000). *Dinamika vyisshikh psihicheskikh funktsii pri obuchenii chteniu i pis'mu pervoklassnikov s zaderzkoi psihicheskogo razvitiya* [The dynamics of higher psychic functions in mentally delayed first graders learning to read and write]. Doctoral dissertation, Moscow City Pedagogical University, 2002. Retrieved February 2008 from http://www.childpsy.ru/dissertations.html

Chiappe, P., & Siegel, L. (1999). Phonological awareness and reading acquisition in English- and Punjabi-speaking Canadian children. *Journal of Educational Psychology, 91*(1), 20-28.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Cohen, J., Cohen, P., West, S., & Aiken, L. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589-608.

Coltheart, M., Rastle, K., Perry C., Langdon, R, & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-254.

Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology, 91*(1), 29-43.

Cossu, G., Shankweiler, D., Liberman, I., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9*, 1-16.

Crombie, M. A. (1997). The effects of specific learning difficulties (dyslexia) on the learning of a foreign language in school. *Dyslexia, 3*, 27-47.

Crystal, D. (1992). *Dictionary of language and languages.* London: Penguin.

Cutting, L, & Scarborough, H. (2006). Prediction of reading comprehension: Relative contribution of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277-299.

de Jong, P. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70,* 75-96.

de Jong, P. (2003). Problems in the acquisition of fluent word decoding in Dutch children. In N. Goulandris (Ed.). *Dyslexia in different languages*. Whurr: London.

de Jong, P., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*(3), 450-476.

de Jong, P., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading, 6*(1), 51-77.

de Jong, P., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology, 95*(1), 22-40.

de Jong, P., & Vrielink, L. (2004). Rapid automatic naming: Easy to measure, hard to improve (quickly). *Annals of Dyslexia, 54*(1), 65-88.

Denckla, M., & Rudel, R. (1976). Rapid automatized naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

Diakidoy, I-A., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing grade levels. *Reading Psychology, 26,* 55-80.

Durgunoglu, A., & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing: An Interdisciplinary Journal, 11*, 281-299.

Ehri, L. (2002). Reading processes, acquisition, and instructional implications. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 167-186). Chichester, UK: Wiley.

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2)*,* 116-125.

Ehri, L. C. (2002). Reading processes, acquisition, and instructional implications. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 167-186). Chichester, UK: Wiley.

Ehri, L., & Wilce, L (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75(1)*, 3-18.

Ehri, L., & Wilce, L (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20*(2), 163-179.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing: An Interdisciplinary Journal, 8,* 453-485.

Elbro, C., Borstrom, I., & Petersen, D. (1998). Predicting dyslexia from kindergarten: The importance of phonological representations of lexical items. *Reading Research Quarterly, 33*(1), 36-60.

Elifson, K., Runyon, R., & Haber, A. (1998). *Fundamentals of social statistics* (3rd ed.). Boston: McGraw-Hill.

Elkonin, D. (1973). USSR. In Dowing, J. (Ed.). *Comparative Reading* (pp. 551-579). New York: Macmillan.

Engle, R., Cantor, J., & Carullo, J (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992.

Everatt, J., Smythe, I., Adams, E., & Ocampo, D. (2000). Dyslexia screening measures and bilingualism. *Dyslexia, 6*, 42-56.

Everatt, J., Smythe, I., Ocampo, D., & Gyarmathy, E. (2004). Issues in the assessment of literacy-related difficulties across language backgrounds: A cross-linguistic comparison. *Journal of Research in Reading, 27*(2), 141-151.

Fawcett, A. (2002). Dyslexia and literacy: Key issues for research. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 11-28). Chichester, UK: Wiley.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Fiez, J. (2000). Sound and meaning: How native language affects reading strategies. *Nature Neuroscience, 3*, 3-5.

Fletcher, M., Coulter, A., Reschly, D., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*(2), 304-331.

Fox, C., & Jones, J. (1998). Uses of Rasch modeling in counseling psychology research. *Journal of Counseling Psychology, 45*(1), 30-45.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301-330). London: Erlbaum.

Frith, U. (2002). Resolving the paradoxes of dyslexia. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 45-68). Chichester, UK: Wiley.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2*, 31-54.

Frost, J., Madsbjerg, S., Niedersoe, J., Olofsson, A., & Sorensen, P. (2005). Semantic and phonological skills in predicting reading development: From 3-16 years of age. *Dyslexia, 11*, 79-92.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*(1), 71-99.

Frost, R., & Bentin, S. (1992). Reading consonants and guessing vowels: Visual word recognition in Hebrew orthography. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 27-44). Amsterdam: North-Holland.

Frost, R., & Kampf, M. (1993). Phonetic recoding of phonologically ambiguous printed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(1), 23-33.

Frost, R., & Katz, L. (1992). *Orthography, phonology, morphology, and meaning.* Amsterdam: North-Holland.

Gang, M., & Siegel, L. (2002). Sound-symbol learning in children with dyslexia. *Journal of Learning Disabilities, 35*(2), 137-157.

Ganschow, L., Sparks, R., & Javorsky, J. (1998). Foreign language learning difficulties: An historical perspective. *Journal of Learning Disabilities, 31*(3), 248-58.

Ganschow, L., Sparks, R., & Schneider, E. (1995). Learning a foreign language: Challenges for students with language learning difficulties. *Dyslexia, 1,* 75-95.

Gathercole, S., & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*(2), 200-213.

Gathercole, S., Baddeley, A., & Willis, C. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology, 82*, 337-346.

Gathercole, S., Hitch, G., Service, E., & Martin, A. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*(6), 966-979.

Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning, 49*(2), 183-217.

Glozman, A. (1999). Russian neuropsychology after Luria. *Neuropsychology Review, 9*(1), 33-44.

Goff, D., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing, 18*, 583-616.

Gonzalez, J., Gonzalez, C., Monzo, A., & Hernandez-Valle, I. (2000). Onset-rime units in visual word recognition in Spanish normal readers and children with reading disabilities. *Learning Disabilities Research and Practice, 15*(3), 135-141.

Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology, 40*, 239-268.

Goswami, U. (1999). Causal connections in beginning reading: The importance of rhyme. *Journal of Research in Reading, 22*(3), 217-240.

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia, 6*, 133-151.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52*, 141-163.

Goswami, U. (2003). Why theories about developmental dyslexia require developmental designs. *Trends in Cognitive Sciences, 7*(12), 534-540.

Goswami, U. (2004, April). Rhyme, rhythm and dyslexia. In T. Hensch (Chair), *Auditory processing and dyslexia.* Symposium conducted at the meeting of the RIKEN BSI and Oxford-Kobe Seminar Joint International Symposium on Dyslexia in Japanese, Chinese, and English, Kobe, Japan.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Erlbaum.

Goswami, U., Ziegler, J., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics, 24*, 235-247.

Gottardo, A. (2002). The relationship between language and reading skills in bilingual Spanish-English speakers. *Topics in Language Disorders*, *22*(5), 46-70.

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7,* 6-10.

Goulandris, N. (2003a). Introduction: Developmental dyslexia, languages and orthographies. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 1-14). London: Whurr.

Goulandris, N. (Ed.). (2003b). *Dyslexia in different languages*: *Cross-linguistic comparisons.* London: Whurr.

Graves, R., & Weinstein, S. (2004). A Rasch analysis of three of the Wisconsin Scales of psychosis proneness: Measurement of schizotypy. *Journal of Applied Measurement, 5*(2), 160-171.

Green, S. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*(3), 499-510.

Green, S., & Salkind, N. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th Ed.). Upper Saddle River, NJ: Pearson.

Grigorenko, E. L. (2003). The difficulty of mastering reading and spelling in Russian. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 92-111). London: Whurr.

Grushevskaya, M. (1982). *Disgraphiya u mladshikh shkol'nikov* [Disgraphiya among elementary school children]. Doctoral dissertation, Moscow City Pedagogical University, 2002. Retrieved February 2008 from http://www.childpsy.ru/dissertations.html

Gyarmathy, E. (2004). Research on dyslexia in Hungarian. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross language comparison and practice guide* (pp. 123-131). Chichester, UK: John Wiley.

Halpern, D. (1999). Review: The Russian perspective on the psychological landscape. *The American Journal of Psychology, 112*(2), 321-324.

Hamilton, W. (1980). *Introduction to Russian phonology and word structure*. Bloomington, IN: Slavica.

Hammill, D. (2004). What we know about correlates of reading. *Council for Exceptional Children, 70*(4), 453-468.

Hansen, J., & Bowey, J. (1994). Phonological analysis skills, verbal working memory, and reading ability in second-grade children. *Child Development, 65*(3), 938-950.

Hatcher, J., & Snowling, M. (2002). The phonological representations hypothesis of dyslexia: From theory to practice. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 69-84). Chichester, UK: Wiley.

Hirose, T., & Hatta, T. (1988). Reading disabilities in modern Japanese children. *Journal of Research in Reading, 11,* 152-160.

Hoien, T., Lundberg, I., Stanovich, K., & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal, 7*, 171-188.

Holopainen, L., Ahonen, T., & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities, 34*(5), 401-414.

Holopainen, L., Ahonen, T., & Lyytinen, H. (2002). The role of reading by analogy in first grade Finnish readers. *Scandinavian Journal of Educational Research, 46*(1), 83-98.

Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127-160.

Hulme, C., Hatcher, P., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*, 2-28.

Inshakova, O. (2004). Problems of dyslexia in Russian. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross-language comparison and practice guide. Part 1: Languages* (pp. 173-178). Chichester, UK: Wiley.

Jenkins, J., Fuchs, L., van den Broek, P., Espin, C., & Deno, S. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729.

Johansson, B. (2006). Cultural and linguistic influence on brain organization for language and possible consequences for dyslexia: A review. *Annals of Dyslexia, 56*(1), 13-50.

Jorm, A., Share, D., MacLean, R., & Matthews, R. (1984). Phonological recoding skills and learning to read: A longidudinal study. *Applied Psycholinguistics, 5,* 201-207.

Joshi, M., & Aaron, P. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology, 21*, 85-97.

Juel, C., Griffith, P., & Gough, P. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*(4), 243-255.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67-84). Amsterdam: North-Holland.

Kaufman, A. (1994*). Intelligent testing with the WISC-3*. New York: Wiley.

Kaufman, A., & Kaufman, N. (2004). *Kaufman Brief Intelligence Test,* 2nd ed. (KBIT-2). Circle Pines, MN: AGS.

Knight, D., & Hynd, G. (2002). The neurobiology of dyslexia. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 29-43). Chichester, UK: Wiley.

Kobayashi, M., Haynes, C., Macaruso, P., Hook, P., & Kato, J. (2005). Effects of mora deletion, nonword repetition, rapid naming, and visual search performance on beginning reading in Japanese. *Annals of Dyslexia, 55*(1), 105-128.

Kornev, A. (1995). *Disleksiya i disgraphiya u detei* [Dyslexia and dysgraphia in children]. St. Petersburg: Gippokrat.

Kornev, A. (2006). Disleksiya i eyo dvoiniki: Kriterii differentsiatsii [Dyslexia and its doubles: Criteria of differentiation]. *Vestnik RAD*, Russian Dyslexia Association. Retrieved January 10, 2007, from http://dyslexia.org.ru/?page=publish/sbornik/korn-01

Kreschenko, Y. (2003). *Psihofiziologicheskie kriterii trudnostei obucheniya pis'mu i chteniu u shkol'nikov mladshikh klassov* [Psychophysiological criteria of difficulty in the learning of reading and writing among elementary school children]. Doctoral dissertation, Moscow City Pedagogical University, 2002. Retrieved March 2008 from http://www.childpsy.ru/dissertations.html

Kubasova, O. (2003). *Literaturnoye chteniye, pervyi klass* [Literature reading, Grade 1]. Smolensk, Russia: 21 Century Association.

Kubasova, O. (2004). *Literaturnoye chteniye, tretyi klass. Chast' 1* [Literature reading, Grade 3. Part 1]. Smolensk, Russia: 21 Century Association.

Kubasova, O. (2005). *Literaturnoye chteniye, vtoroi klass. Chast' 1* [Literature reading, Grade 2. Part 1]. Smolensk, Russia: 21 Century Association.

Lalaeva, R. (2004). *Narusheniya chteniya i pis'ma u mladshih shkol'nikov* [Reading and writing disabilities in elementary school children]. Rostov, Russia: Fenix.

Landerl, K. (2001). Word recognition deficits in German: More evidence from a representative sample. *Dyslexia, 7,* 183-196.

Landerl, K. (2003). Dyslexia in German-speaking children. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 15-32). London: Whurr.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia. A German-English comparison. *Cognition, 63*, 315-334.

Leong, C. K. (1999). Phonological and morphological processing in adult students with learning/reading disabilities. *Journal of Learning Disabilities, 32*(3), 224-238.

Lie, A. (1991). Effects of a training program for stimulating skills in word analysis in first-grade children. *Reading Research Quarterly, 26*(3), 234-250.

Linacre, J. M. (1998). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement, 2,* 266-283.

Linacre, J. M. (2006). Winsteps (Version 3.64.2) [Computer Software]. Chicago: Winsteps.com.

Linacre, M., & Wright, B. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370. Retrieved March 17, 2005, from http://www.rasch.org/rmt/rmt83.htm

Lukatela, G., & Turvey, M. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General, 123*(2), 107-128.

Lundberg, I., Frost, J., & Petersen, O-P. (1988). Effects of an extensive program to stimulate phonological awareness in preschool children. *Reading Research Quarterly, 23*(3), 263-284.

Lunz, M., Stahl, J., & Wright, B. (1996). The invariance of judge severity calibrations. In G. Engelhard & M. Wilson (Eds*). Objective measurement: Theory into practice, Volume 3* (pp. 99-112). Norwood, NJ: Ablex.

Lyon, G. (1996). Learning disabilities. *The Future of Children, 6*(1), 54-76.

Lyytinen, H., Aro, M., & Holopainen, L. (2004). Dyslexia in highly orthographically regular Finnish. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross language comparison and practice guide* (pp. 81-91). Chichester, UK: John Wiley.

Manis, F., Seidenberg, M., & Doi, L. (1999). See Dick RAN: Rapid naming and the longitudinal prediction of reading subskills in first and second graders. *Scientific Studies of Reading, 3*(2), 129-157.

Mann, V. A., & Foy, J. G. (2003). Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia, 53*, 149-173.

Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities*, *17*, 592-99.

Mannai, H., & Everatt, J. (2005). Phonological processing skills as predictors of literacy amongst Arabic speaking Bahraini children. *Dyslexia, 11*, 269-291.

Marshall, C. M., Snowling, M. J., & Bailey, P. J. (2001). Rapid auditory processing and phonological processing in dyslexic and normal readers. *Journal of Speech, Hearing, and Language Research, 44*, 925-940.

Martin, F., Claydon, E., Morton, A., Binns, S., & Pratt, C. (2003). The development of orthographic and phonological strategies for the decoding of words in children. *Journal of Research in Reading, 26*(2), 191-204.

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.

McBride-Chang, C., & Kail, R. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73*(5), 1392-1407.

McBride-Chang, C., & Manis, F. (1996). Structural invariance in the associations of naming speed, phonological awareness, and verbal reasoning in good and poor readers: A test of the double deficit hypothesis. *Reading and Writing: An Interdisciplinary Journal, 8*, 323-339.

Megherbi, H., Seigneuric, A., & Ehrlich, M-F. (2006). Reading comprehension in French 1st and 2nd grade children: Contribution of decoding and language comprehension. *European Journal of Psychology of Education, 21*(2), 135-147.

Meyer, M., Wood, F., Hart, L., & Felton, R. (1998). Selective predictive value of rapid automatized naming in poor readers. *Journal of Learning Disabilities, 31*(2), 106-117.

Miller, D., & Kunce, J. (1973). Prediction and statistical overkill revisited. *Measurement and Evaluation in Guidance, 6*(3), 157-163.

Milton, S. (1986). A sample size formula for multiple regression studies. *The Public Opinion Quarterly*, *50*(1), 112-118.

Morrison, C. (1996). Predicting academic performance in college: An investigation of the utility of the graded response. In G. Engelhard & M. Wilson (Eds*). Objective measurement: Theory into practice, Volume 3* (pp. 45-64). Norwood, NJ: Ablex.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*(3), 320-337.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 65*, 370-396.

Naglieri, J. (1996). An examination of the relationship between intelligence and reading achievement using the MAT–SF and MAST. *Journal of Psychoeducational Assessment, 14*, 65-69.

Naglieri, J. (2001). Do ability and reading achievement correlate? *Journal of Learning Disabilities, 34*(4), 304-305.

Naglieri, J., & Ronning, M. (2000). The relationships between general ability using the NNAT and SAT reading achievement. *Journal of Psychoeducational Assessment, 18*, 230-239.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*(2), 154-167.

Nation, K., & Snowling, M. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39*, 85-101.

Nation, K., & Snowling, M. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*, B1-B13.

Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356.

Nelson, H. E., & Warrington, E. K. (1980). An investigation of memory functions in dyslexic children. *British Journal of Psychology, 71*, 487-503.

Nikolopoulos, D., Goulandris, N., & Snowling, M. (2003). Developmental dyslexia in Greek. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 53-67). London: Whurr.

Ogarkina, A. (2002) *Vuiyavlenie predraspolozennosti k disleksii u doshkol'nikov.* [Identifying children at risk for dyslexia among preschoolers]. Doctoral dissertation, Moscow City Pedagogical University, 2002. Retrieved April 2008 from http://www.childpsy.ru/dissertations.html

Oktay, A., & Aktan, E. (2002). A cross-linguistic comparison of phonological awareness and word recognition in Turkish and English. *International Journal of Early Years Education, 10*(1), 37-48.

Olofsson, A. (2003). The dyslexic reader and the Swedish language. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 137-156). London: Whurr.

Oney, B., & Goldman, S. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme-phoneme correspondences. *Journal of Educational Psychology, 76*(4), 557-568.

Ozhegov, C., & Shvedova, H. (2003). *Tolkovyi slovar' russkogo yazyika, 4-e izdanie* [Definitional dictionary of the Russian language, 4th edition]. Moscow: ITI Technology.

Paap, K., Noel, R., & Johansen, L. (1992). Dual-route models of print to sound: Red herrings and real horses. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 293-318). Amsterdam: North-Holland.

Passenger, T., Stuart, M., & Terrell, C. (2000). Phonological processing and early literacy. *Journal of Research in Reading, 23*, 55-66.

Paulesu, E., Demonet, J-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001, March 16). Dyslexia: Cultural diversity and biological unity. *Science, 291,* 2165-2167.

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R., et al. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain, 119,* 143-157.

Paulesu, E., McCrory, E., Fazio, F., Memoncello, N., Brunswick, N., Cappa, S., et al. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*, 91-96.

Pennington, B. F., Van Orden, G., Smith, S., Green, P., & Haith, M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Development, 61*, 1753-78.

Perfetti, C. (2003). The universal grammar of reading. *Scientific Studies of Reading, 7*(1), 3-24.

Perfetti, C., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633-643.

Perfetti, C., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24-33.

Perfetti, C., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill Palmer Quarterly, 33*, 283-319.

Perfetti, C., Bell, L., & Delaney, S. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27*, 59-70.

Philimonenko, Y., & Timofeev, V. (2001). *Wechsler Intelligence Scale for Children* (Russian version). St. Petersburg: Imaton.

Pikulski, J., & Chard, D. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher, 58*(6), 510-519.

Porpodas, C. (1999). Patterns of phonological and memory processing in beginning readers and spellers of Greek. *Journal of Learning Disabilities, 32*(5), 406-416.

Pugh, K., Sandak, R., Frost, S., Moore, D., & Mencl, E. (2005). Examining reading development and reading disability in English language learners: Potential contributions from functional neuroimaging. *Learning Disabilities Research and Practice, 20*(1), 24-30.

Rack, J., Snowling, M., & Olson, R. (1992). The nonword reading deficit in dyslexia: A review. *Reading Research Quarterly, 27*, 29-53.

Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841-865.

Ramzaeva, T. (2005a). *Russkii yazyik, vtoroi klass. Chast' 1* [Russian language, Grade 2. Part 1]. Moscow: Drofa.

Ramzaeva, T. (2005b). *Russkii yazyik, vtoroi klass. Chast' 2* [Russian language, Grade 2. Part 2]. Moscow: Drofa.

Ratanova, T. (2005). *Diagnostika umstvennih sposobnostei detei* [Diagnosis of intellectual abilities in children]. Moscow: Flinta.

Reitsma, P. (2004). Developmental dyslexia in the Dutch language. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross language comparison and practice guide* (pp. 53-61). Chichester, UK: John Wiley.

Robertson, J. (2000). The neuropsychology of modern foreign language learning. In L. Peer & G. Reid (Eds.), *Multilingualism, literacy and dyslexia: A challenge for educators* (pp. 203-210). London: David Fulton.

Rupley, W., Willson, V., & Nichols, W. (1998). Exploration of the developmental components contributing to elementary school children's reading comprehension. *Scientific Studies of Reading, 2*(2), 143-158.

Russian Ministry of Education. (2000). *Gosudarstvennui obrazovatel'nyi standart vischego professional'nogo obrazovaniya. Spetsial'nost' 020400—psihologiya. Kvalifikatsiya—psiholog, prepodavatel' psihologii.* [State educational standards of higher professional education. Specialty 020400, psychology. Qualification: psychologist, teacher of psychology]. Moscow: Russian Ministry of Education.

Savage, R. (2001). The 'Simple View' of reading: Some evidence and possible implications. *Educational Psychology in Practice, 17*(1), 17-33.

Savage, R. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading, 10*(2), 143-164.

Sawyer, R. (1982). Sample size and the accuracy of predictions made from multiple regression equations. *Journal of Educational Statistics, 7*(2), 91-104.

Scarborough, H. (1990). Very early language deficits in dyslexic children. *Child Development, 61*(6), 1728-1743.

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*, 439-449.

Schiermeier, Q. (2007, October 4). The battle for Russia's brains. *Nature, 449*(7162), 524-527.

Schneider, W., Kuspert, P., Roth, E., & Vise, M. (1997). Short- and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology, 66*, 311-340.

Seidenberg, M. (1992). Beyond orthographic depth in reading: Equitable division of labor. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 85-118). Amsterdam: North-Holland.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143-175.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K., Fletcher, J., Brady, S., et al. (1999). Comprehension and decoding patterns of association in children with reading difficulties. *Scientific Studies of Reading, 39(*1), 69-94.

Share, D. (1995). Phonological recoding and self-teaching: *Sine qua non* of reading acquisition. *Cognition, 55*, 151-218.

Share, D. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95-129.

Share, D. (2003). Dyslexia in Hebrew. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 208-234). London: Whurr.

Share, D., & Leikin, M. (2004). Developmental dyslexia in the Hebrew language. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross language comparison and practice guide* (pp. 113-120). Chichester, UK: John Wiley.

Sheridan, B., & Puhl, L. (1996). Evaluating an indirect measure of student literacy competencies in higher education using Rasch measurement. In G. Engelhard & M. Wilson (Eds*). Objective measurement: Theory into practice, Volume 3* (pp. 19-44). Norwood, NJ: Ablex.

Singleton, C. (2002). Dyslexia: Cognitive factors and implications for literacy. In G. Reid & J. Wearmouth (Eds.), *Dyslexia and literacy: Theory and practice* (pp. 115-130). Chichester, UK: Wiley.

Smith, E. (2004). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. In E. Smith & R. Smith (Eds.). *Introduction to Rasch measurement: Theory, models, and applications* (pp. 575-600). Maple Grove, MN: JAM Press.

Smith, R. (1991). The distributional properties of Rasch item fit statistics. *Educational and Psychological Measurement, 51*, 541-565.

Smith, R. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modeling, 3*, 25-40.

Smith, R. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement 1*(2), 199-218.

Smith, R. (2003). *Rasch measurement models: Interpreting WINSTEPS and FACETS output*. Maple Grove, MN: JAM Press.

Smith, R. (2004). Fit analysis in latent trait measurement models. In E. Smith & R. Smith (Eds.). *Introduction to Rasch measurement: Theory, models, and applications.* (pp. 73-92). Maple Grove, MN: JAM Press.

Smith, R., & Miao, C. (1994). Assessing unidimensionality for Rasch measurement. In M. Wilson (Ed.). *Objective measurement: Theory into practice, Volume 2* (pp. 316-327). Norwood, NJ: Ablex.

Smythe, I., & Everatt, J. (2000). Dyslexia diagnosis in different languages. In L. Peer & G. Reid (Eds.), *Multilingualism, literacy and dyslexia: A challenge for educators* (pp. 12-21). London: David Fulton.

Smythe, I., Everatt, J., & Salter, R. (Eds.). (2004). *International book of dyslexia: A cross-language comparison and practice guide.* Chichester, UK: Wiley.

Snowling, M. (2000). *Dyslexia.* Oxford, UK: Blackwell.

Snowling, M., Hulme, C., & Goulandris, N. (1994). Word recognition in developmental dyslexia: A connectionist interpretation. *Quarterly Journal of Experimental Psychology, 47A*, 895-916.

Snowling, M., Stackhouse, J., & Rack, J. (1986). Phonological dyslexia and dysgraphia: A developmental analysis. *Cognitive Neuropsychology, 3*, 309-333.

Sprenger-Charolles, L., Colé, P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia: Evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology*, *54*(2), 87-104.

Stahl, S., & Murray, B. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*(2), 221-234.

Stanovich, K. E. (1981). Relationships between word decoding speed, general name-retrieval ability, and reading progress in first-grade children. *Journal of Educational Psychology, 73*(6), 809-815.

Stanovich, K. E. (1991). Discrepancy definition of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26*, 7-29.

Stanovich, K. E. (2000). *Progress in understanding reading*. New York: Guilford Press.

Stanovich, K., Cunningham, A., & Freeman, D. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19(3)*, 278-303.

Swan, D., & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology, 66*, 18-41.

Swanson, H. (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement of children and adults with learning disabilities? *Journal of Learning Disabilities, 27*(1), 34-50.

Swanson, H., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence, 21*, 83-108.

Swanson, H., Trainin., G., Necoechea, D., & Hammill, D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*(4), 407-440.

Szczerbinski, M. (2003). Dyslexia in Polish. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 68-91). London: Whurr.

Tabachnick, B., & Fidell, L. (1989). *Using multivariate statistics*. (2nd ed.). New York: Harper & Row.

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics*. (4th ed.). Boston: Allyn & Bacon.

Thurlow, M., Liu, K., Albus, D., & Shyyan, V. (2003). *Literature on early instruction in four languages* (Project No. ED-01-PO-3927). Minneapolis: University of Minnesota, National Center on Educational Outcomes. Retrieved September 11, 2007, from http://www.ncela.gwu.edu/ resabout/research/literature_ early_literacy.pdf

Time for a fresh start. (2007, October 4). *Nature, 449*(7162), 507.

Tiu, R., Thompson, L., & Lewis, B. (2003). The role of IQ in a component model of reading. *Journal of Learning Disabilities, 36*(5), 424-436.

Torgesen, J., Wagner, R., & Rashotte, C. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27*(5), 276-286.

Torgesen, J., Wagner, R., Rashotte, C., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid naming ability to the growth of word-reading skills in second-to-fifth-grade children. *Scientific Studies of Reading, 1*(2), 161-185.

Tractenberg, R. E. (2002). Exploring hypotheses about phonological awareness, memory, and reading achievement. *Journal of Learning Disabilities, 35*(5), 407-425.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology, 39*, 161-181.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language, 25,* 476-491.

Tsvetkova, L. (1997). *Neuropsihologiya scheta, pis'ma i chteniya: Narusheniye i vosstanovleniye* [Neuropsychology of counting, writing, and reading: Disorders and remediation]. Moscow: Yurist. Retrieved March 17, 2008, from http://pedlib.ru/Books/2/0155/index.html

Tunmer, W., & Hoover, W. (1993). Phonological recoding skill and beginning reading. *Reading and Writing: An Interdisciplinary Journal, 5,* 161-179.

Tunmer, W., Herriman, M., & Nesdale, A (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly, 23(3)*, 134-158.

Tupper, D. (1999). Introduction: Alexander Luria's continuing influence on worldwide neuropsychology. *Neuropsychology Review, 9*(1), 1-7.

Turner, M. (1997). *Psychological assessment of dyslexia.* London: Whurr.

Turner, M., & Ridsdale, J. (2004). *The digit memory test.* Retreived July 15, 2005, from http://www.dyslexia-inst.org.uk/pdffiles/Digit.pdf

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional word reading via phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(2), 336-349.

Van den Bos, K. (1998). IQ, phonological awareness and continuous-naming speed related to Dutch poor decoding children's performance on two word identification tests. *Dyslexia, 4*, 73-89.

Van Etten, G. (1978). A look at reading comprehension. *Journal of Learning Disabilities, 11*(1), 42-46.

Van Ijzendoorn, M., & Bus, A. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly, 29*(3), 266-275.

Van Orden, G., Johnston, J., & Hale, B. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 371-386.

Velichenkova, O. (2002). Spetsificheskiye narusheniya pis'ma u detei [Specific writing disorders in children]. Doctoral dissertation, Moscow City Pedagogical University, 2002. Retrieved February 2008 from http://www.childpsy.ru/dissertations.html

Vellutino, F., Fletcher, J., Snowling, M., & Scanlon, D. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2-40.

Wade-Wooley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning, 49*(3), 447-471.

Wagner, R., & Torgesen, J. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192-212.

Wagner, R., Torgesen, J., Rashotte, C., Hecht, S., Barker, T., Burgess, S., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33*(3), 468-479.

What the scientists say. (2007, October 4). *Nature, 449*(7162), 528-531.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics, 14*, 1-34.

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: Evidence from children learning to read German. *Journal of Experimental Child Psychology, 61*, 80-90.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51,* 91-103.

Wimmer, H., & Hummer, P. (1990). How German speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11,* 349-368.

Wimmer, H., Landerl, K., & Schneider, W. (1994). The role of rime awareness in learning to read a regular orthography. *British Journal of Developmental Psychology, 12*, 469-484.

Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: A deficit in skill-automatization or a phonological deficit? *Scientific Studies of Reading, 2*(4), 321-340.

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92*(4), 668-680.

Wolf, M., & Bowers, P. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415-438.

Wolf, M., Bowers, P., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33*(4), 387-407.

Wright, B., & Stone, M. (1979). *Best test design: Rasch measurement.* Chicago: MESA.

Wydell, T. (2003). Dyslexia in Japanese and the hypothesis of granularity and transparency. In N. Goulandris (Ed.), *Dyslexia in different languages: Cross-linguistic comparisons* (pp. 255-276). London: Whurr.

Wydell, T., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition, 70*, 273-305.

Wydell, T., & Kondo, T. (2003). Phonological deficit and the reliance on orthographic approximation for reading: A follow-up study on an English-Japanese bilingual with monolingual dyslexia. *Journal of Research in Reading*, *26*(1), 33-48.

Yamada, J. (2004). Learning to read in Japanese: Word to sentence reading. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia: A cross-language comparison and practice guide. Part 1: Languages* (pp. 139-146). Chichester, UK: Wiley.

Yamada, J., & Banks, A. (1994). Evidence for and characteristics of dyslexia among Japanese children. *Annals of Dyslexia, 44,* 105-19.

Yopp, H. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*(2), 159-177.

Zelenina, L., & Khokhlova, T. (2004). *Russkii yazyik, pervyi klass* [Russian language, Grade 1]. Moscow: Prosveshenie.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across orthographies: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3-29.

Ziegler, J., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science, 9*(5), 429-453.

Ziegler, J., Perry, C., Jacobs, A., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12(*5), 379-384.

Ziegler, J., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Korne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*, 169-193.

APPENDICES

APPENDIX A

RHYME DETECTION

Examiner: Listen as I say three words. Two of them sound similar, but the third one does not. Tell me which one sounds different from the other two.

Example: мой, бой, вор
Practice item: сок, рок, луч

Test items:

1. год, рот, сон
2. тень, щель, день
3. час, шар, жар
4. мыл, пил, бил
5. пол, кол, мел
6. дуб, лоб, зуб
7. слог, блок, флот
8. бровь, срок, кровь
9. брат, град, мрак
10. друг, круг, брак
11. дверь, хворь, зверь
12. клин, блин, слон
13. бинт, винт, зонт
14. порт, март, сорт
15. весть, кость, гость
16. лоск, воск, пост
17. мост, воск, рост
18. риск, писк, лист

APPENDIX B

INITIAL ODDITY

Examiner: Listen as I say three words. Two of them start with the same sound but the third one starts with a different one. Tell me which one starts with a different sound.

Example: мой, мак, луч
Practice item: сок, рот, сын

Test items:

1. лось, лук, сад
2. гусь, пар, пот
3. рак, рысь, лев
4. сор, щит, сын,
5. шар, жир, жук
6. снег, сват, звон
7. внук, враг, флот
8. стул, брат, слон
9. звук, клён, край
10. класс, круг, двор

APPENDIX C

FINAL ODDITY

Examiner: Listen as I say three words. Two of them end with the same sound but the third one ends with a different one. Tell me which word ends with a different sound.

Example: сын, пан, дул
Practice item: быть, даль, мать

Test items:

1. сок, лук, вар
2. час, рис, конь
3. быт, мел, вол
4. пот, зуб, щит
5. лев, ров, даль
6. мост, жизнь, бант
7. цирк, воск, лесть
8. сорт, пуск, бинт
9. бунт, парк, лифт
10. март, риск, полк

APPENDIX D

PHONEME SEGMENTATION

Examiner: I am going to say some made-up words. Listen and when I finish saying a word, tap out the number of sounds in it.

Example: рун
Practice item: кип

Test items:

1. сик	4. крун	7. бринт	10. штрум	13. сприкс
2. моть	5. слош	8. цверк	11. стран	
3. чаф	6. внак	9. псольт	12. всполь	

APPENDIX E

PHONEME DELETION

Examiner: I'm going to say some made-up words and I want you to drop a particular sound from each word—I'll tell you which one. And then I want you to say the new word.

Example: рёш (without "р")
Practice item: шлось (without "ш")

Test items:

1. мон without "м"
2. рум without "р"
3. книс without "к"
4. плёт without "п"
5. брок without "р"
6. твень without "в"
7. чакс without "к"
8. нольчь without "ль"
9. даф without "ф"
10. тым without "м"
11. гаст without "т"
12. мирк without "к"

APPENDIX F

NONWORD REPETITION

Examiner: I'm going to say some made-up words. Listen carefully and after I finish saying a word, repeat it as best as you can.

Example: мун
Practice item: лум-сип

Test items:

1. тул	7. ру-бит	13. бар-ка-зон	19. па-но-ри-тан
2. сеп	8. пен-тел	14. бан-ди-фер	20. ти-ну-ра-зот
3. бамс	9. бал-коп	15. сти-кэ-рин	21. ам-плу-вар-синт
4. грал	10. прин-дол	16. ски-ти-культ	22. чер-кли-сте-ронк
5. хонт	11. таф-лест	17. кра-сте-рер	23. фон-тран-са-мист
6. клирт	12. сма-пент	18. гли-сте-ринк	24. клон-зер-ска-польт

APPENDIX G

FORWARD DIGIT SPAN

Examiner: Listen carefully as I say some numbers. They will get longer and longer. Repeat each number as best as you can.

Example: 35 (pronounced “three, five”)
Practice item: 2, 5, 6 (two, five, six)

Test items:

Trial 1	Trial 2
43	16
792	847
5941	7253
93872	75396
152649	216748
3745261	4925316
82973546	69174253
246937185	371625948

APPENDIX H

RAN-OBJECT

Examiner (after confirming that the child is familiar with the names of the objects): I want you to name these objects as fast as you can, OK? You can start now.

APPENDIX I

RAN-DIGIT

Examiner (after confirming that the child is familiar with the names of the digits): I want you to name these digits as fast as you can, OK? You can start now.

2	3	5	6	7
5	2	3	7	6
5	6	7	3	2
7	3	6	5	2
3	2	5	6	7
6	7	2	5	3
3	5	7	2	6
5	3	6	2	7
2	3	6	7	5
7	5	3	2	6

APPENDIX J

WORD READING

Examiner: I want you to read these words as fast as you can and without mistakes.

List 1

1. лес	2. пень	3. крот	4. волк
5. тигр	6. весна	7. доска	8. идёт
9. якорь	10. друзья	11. серьёзно	12. поздравил
13. сплоховал	14. множество	15. подснежник	16. неизвестно
17. спрашивает	18. прожужжала	19. заблудилась	20. подглядывать

List 2

1. бой	2. сон	3. дуть	4. мчал
5. корм	6. шуба	7. дело	8. ссора
9. голос	10. чувство		

List 3

1. разбойник	2. спросонья	3. надувной	4. помчался
5. кормушка	6. полушубок	7. безделушка	8. перессорить
9. голосистый	10. почувствовал		

List 4

1. коридор	2. пластилин	3. превратить	4. отшвырнул
5. встревожить	6. пожалуйста	7. отечество	8. разъярился
9. действительно	10. всполошились		

APPENDIX K

NONWORD READING

Examiner: I want you to read these words as fast as you can and without mistakes. These are not real words but you try to read them as best as you can.

List 1 (nonwords visually dissimilar from real words)

1. сой	2. бон
3. корс	4. снем
5. стур	6. дроль
7. башай	8. калей
9. кисан	10. чигра
11. леньтю	12. дышлан
13. безарё	14. изкабуш
15. прунажи	16. кадашран
17. путыйшис	18. буникдиль
19. пенерифул	20. финоризер
21. контрапонист	22. перплистеронк
23. эмплифорвент	24. блонтерстапинг

List 2 (nonwords visually similar to real words)

1. лас	2. наж
3. кень	4. дыль
5. тугр	6. сталб
7. авощ	8. вьёга
9. завыд	10. юскра
11. зувисть	12. встроча
13. магазон	14. конфата
15. коруткий	16. дровосяк
17. спрытался	18. подпрагнул
19. волныние	20. удивалась
21. эскалутор	22. раскалянный
23. предупредал	24. тропуческий

APPENDIX L

READING COMPREHENSION

Examiner: Now let's read. I want you to read these stories aloud and then I will ask you some questions about each story. Answer my questions as fully as you can.

Story 1

В доме у нас жил ёжик. Он был ручной. Когда его гладили, он прижимал к спине колючки и делался совсем мягким. За это мы его прозвали Пушком.

Вопросы:

1. О ком говорится в этом рассказе?
2. Почему ежа прозвали Пушком?

Story 2

В детском саду было много игрушек. Ребята играли все вместе. Им было весело. Только Дима не играл. Он собрал около себя целую кучу игрушек и охранял их от детей. "Моё! Моё!" – кричал он, закрывая игрушки руками. Ему было скучно.

Вопросы:

1. Что делали ребята в детском саду?
2. Почему Дима не играл?
3. Почему ему было скучно?

Story 3

Мальчик играл и разбил нечаянно дорогую чашку. Никто не видал.
Отец пришел и спросил: "Кто разбил?" Мальчик затрясся от страха и сказал: "Я". Отец сказал: "Спасибо, что правду сказал. Правда всего дороже".

Вопросы:

1. Что сделал мальчик?
2. Почему он затрясся от страха на вопрос отца?
3. Почему отец сказал спасибо?

Story 4

Однажды к избе лесника из чащи прибежал барсук. На голове у него стеклянная банка надета, как скафандр. Смотрит барсук сквозь банку, а снять ее не может. Что-то вкусное было в банке, вот зверёк и попался. За помощью он прибежал к человеку. Лесник аккуратно разбил банку и освободил беднягу. Счастливый барсук быстро скрылся в ельнике.

Вопросы:

1. Что заставило барсука обратиться за помощью к человеку?
2. Как голова барсука оказалась в банке?
3. Куда прибежал зверёк?
4. Как лесник помог бедняге?

Story 5

Самая маленькая птичка – колибри. Она размером с бабочку. Самое большое животное на Земле – слон. Он достигает высоты до трёх метров. Крошечная птичка и огромный слон живут в тёплых странах.

Вопросы:

1. Какая птичка самая маленькая?
2. Какого она размера?
3. Какой высоты достигает слон?
4. Где живут крошечеая птичка и огромный слон?

APPENDIX M

LISTENING COMPREHENSION

Examiner: I'm going to read you some stories and then I'll ask you some questions about each story. Listen and try to answer the questions as fully as you can.

Story 1

Мы с папой гуляли в парке. Вдруг на дорожку выскочил пушистый зверек. Это была ручная белка. Я бросил ей орех. Она схватила его и быстро зарыла под сосной. Так белочка делала свои запасы.

Вопросы:

1. Где произошел этот случай?
2. Зачем белка зарыла орех?

Story 2

Кто такие пингвины? Пингвины – это птицы. Но птицы особенные. Они не умеют летать. Пингвины хорошо плавают. Пищу они добывают в воде, а большую часть суток проводят на суше.

Вопросы:

1. Кто такие пингвины?
2. В чем особенность этих птиц?
3. Где они добывают себе пищу?

Story 3

Люди стали говорить очень давно. Им нужно было добывать пищу, строить жилища. Сообща это делать проще. И люди звали друг друга на помощь. Так появилась устная речь.

Вопросы:

1. О чём говорится в этом рассказе?
2. Как появилась устная речь?
3. Для чего люди звали друг друга на помощь?

Story 4

Витя потерял завтрак. На большой перемене все ребята завтракали, а Витя стоял в сторонке. “Почему ты не ешь?” – спросил его Коля. “Я завтрак потерял”. “Плохо”, - сказал Коля, откусывая большой кусок белого хлеба. “А где ты его потерял?” – спросил Миша. “Не знаю”, – тихо сказал Витя и отвернулся.”Ты, наверное, в кармане нёс, а надо в сумку класть”, - сказал Миша. А Володя ничего не спросил. Он подошел к Вите, разломил пополам кусок хлеба с маслом и протянул товарищу: “Бери, ешь!”

Вопросы:

1. Где произошел этот случай?
2. Почему Витя стоял в сторонке и не завтракал?
3. Сколько товарищей подошли к нему?
4. Помог ли кто-нибудь из них Вите?

Story 5

Серёжа получил письмо от друга. Он живёт в Москве. Андрей пишет о своём классе. У него хорошие друзья. Недавно у него был день рождения, и ему подарили коллекцию марок.

Вопросы:

1. Кто получил письмо от друга?
2. Где живёт Андрей?
3. О чём он пишет в письме?
4. Что ему подарили на день рождения?

APPENDIX N

READING COMPREHENSION (ENGLISH TRANSLATION)

Story 1

In our home, there lived a hedgehog. When it was patted, it lowered its quills and became completely soft. Because of that, we called it Fluffy.

Questions:

1. Who is the main character of this story? [Lit.: Who is this story about?]
2. Why was the hedgehog called Fluffy?

Story 2

There were many toys in the kindergarten. The children played all together. They were having fun. Only Dima wasn't playing. He gathered a whole bunch of toys to himself and guarded them from the children. "They're mine, mine," he cried, covering them with his hands. He was bored.

Questions:

1. What were the children doing in the kindergarten?
2. Why wasn't Dima playing?
3. Why was he bored?

Story 3

The boy was playing and he accidentally broke an expensive cup. No one saw it happen. His father came and asked: "Who broke the cup?" The boy started to tremble with fear and said, "I did." The father said, "Thank you for telling the truth. Truth is the most precious thing."

Questions:

1. What did the boy do?
2. Why did he start trembling with fear when he heard his father's question?
3. Why did the father say "thank you"?

Story 4

One day there came running from the depth of the forest to the forester's hut a badger. On its head, like a spacesuit helmet is a glass jar. The badger is looking through the jar, but cannot take it off. There was something tasty in the jar, and that's how the little animal got trapped. For help, it came running to a human. The forester gently broke the jar and freed the poor thing. The happy badger disappeared in the fur-grove.

Questions:

1. What made the badger seek help from a human?
2. How did its head end up inside the jar?
3. Where did the little animal come running to?
4. How did the forester help the poor thing?

Story 5

The smallest bird is the humming bird. It's the size of a butterfly. The biggest animal on Earth is the elephant. It can reach the height of three meters. The tiny bird and the giant elephant live in warm countries.

Questions:

1. What bird is the smallest?
2. What size is it?
3. What height can the elephant reach?
4. Where do the tiny bird and the giant elephant live?

APPENDIX O

LISTENING COMPREHENSION (ENGLISH TRANSLATION)

Story 1

My father and I were strolling in the park. Suddenly a fluffy animal ran out onto the path. It was a tame squirrel. I threw it a nut. The squirrel grabbed it and quickly buried it under a pine tree. That's how the little squirrel was making its supply of food.

Questions:

1. Where did this incident take place?
2. Why did the squirrel hide the nut?

Story 2

What are penguins? Penguins are birds. But [they are] a special kind of birds. They cannot fly. Penguins swim well. They find food in the water and spend most of the day on land.

Questions:

1. What are penguins?
2. How are they different from other birds?
3. Where do they find food?

Story 3

People began to talk a very long time ago. They needed to find food, build dwellings. This is easier to do through cooperation. So people called on each other for help. And that's how oral speech appeared.

Questions:

1. What is this story about?
2. How did oral speech appear?
3. Why did people call on each other for help?

Story 4

Vitya lost his breakfast. During the long recess, all the children were eating breakfast, but Vitya was standing apart from them. “Why aren’t you eating?” asked Kolya. “I lost my breakfast.” “That’s too bad,” said Kolya, biting off a big piece of white bread. “Where did you lose it?” asked Misha. “I don’t know,” Vitya said softly and turned away. “You must have carried it in your pocket, and you should put it in your bag,” said Misha. And Volodya didn’t say anything. He came up to Vitya, broke in half a piece of buttered bread, and handed one half over to his friend. “Here, eat!”

Questions:

1. Where did the story happen?
2. Why was Vitya standing apart from the other children and not eating?
3. How many friends approached him?
4. Did any of them help him?

Story 5

Seryozha received a letter from his friend. He lives in Moscow. Andrei writes about his class. He’s got good friends. Recently he had a birthday and was given a collection of stamps for a present.

Questions:

1. Who received a letter from a friend?
2. Where does Andrei live?
3. What is he writing about in the letter?
4. What did he get for his birthday?

APPENDIX P

RUSSIAN TEST BATTERY

Фамилия ученика, возраст_____

Блок 1: Фонологические Тесты

Фонологический блок состоит из пяти тестов. Эти тесты даются ребенку в том порядке, в каком они указаны. Порядок заданий в каждом из этих тестов произвольный и выбирается экзаменатором.

Тест 1 (rhyme detection)

Экзаменатор: «Я тебе сейчас произнесу три слова. Два из них звучат похоже, а третье—нет. Ты их послушай и скажи, какое слово звучит по другому, то есть отличается от двух других. Вот например....» (правильный ответ выделен).

Пример и проба: мой, бой, **вор**; **луч,** сок, рок

1. год, рот, **сон**	7. слог, блок, **флот**	13. бинт, винт, **зонт**
2. тень, **щель,** день	8. бровь, **срок,** кровь	14. порт, **март,** сорт
3. **час,** шар, жар	9. брат, град, **мрак**	15. **весть,** кость, гость
4. **мыл,** пил, бил	10. друг, круг, **брак**	16. лоск, воск, **пост**
5. пол, кол, **мел**	11. дверь, **хворь,** зверь	17. мост, **воск,** рост
6. дуб, **лоб,** зуб	12. клин, блин, **слон**	18. риск, писк, **лист**

Тест 2 (initial oddity)

Экзаменатор: "Я тебе сейчас произнесу три слова. Два из них начинаются с одного и того же звука, а третье—нет. Ты их послушай и скажи мне, какое слово начинается с другого звука. Вот например...." (правильный ответ выделен).

Пример и проба: мой, мак, **луч;** сок, **рот**, сын

1. лось, лук, **сад**
2. гусь, **пар,** пот
3. рак, рысь, **лев**
4. сор, **щит,** сын,
5. **шар**, жир, жук
6. снег, сват, **звон**
7. внук, враг, **флот**
8. стул, **брат**, слон
9. **звук**, клён, край
10. класс, круг, **двор**

Тест 3 (final oddity)

Экзаменатор: «Я тебе сейчас произнесу три слова. Два из них кончаются на один и тот же звук, а третье—нет. Ты их послушай и скажи мне, какое слово заканчивается на другой звук. Вот например....» (правильный ответ выделен).

Пример и проба: сын, пан, **дул**; быть, **даль,** мать

1. сок, лук, **вар**
2. час, рис, **конь**
3. **быт**, мел, вол
4. пот, **зуб,** щит
5. лев, ров, **даль**
6. мост, **жизнь,** бант
7. цирк, воск, **лесть**
8. сорт, **пуск**, бинт
9. бунт, **парк,** лифт
10. **март**, риск, полк

Тест 4 (phoneme segmentation)

Экзаменатор: «Я тебе сейчас скажу слово, только оно не настоящее, а придуманное. Ты его послушай и постучи карандашом по столу столько раз, сколько в этом слове звуков. Вот, напирмер, если ты думаешь, что в слове 3 звука, постучи по столу 3 раза, а если 5 – то 5 раз. Вот если я тебе скажу слово "рун", ты постучишь по столу 3 раза, потому что в этом слове 3 звука—р, у, н. Ну, давай попробуем» (дать пробное слово).

Пример и проба: Рун (3 раза); **кип** (три раза)

1. сик	4. крун	7. бринт	10. штрум	13. сприкс
2. моть	5. слош	8. цверк	11. стран	
3. чаф	6. внак	9. псольт	12. всполь	

Тест 5 (phoneme deletion)

Экзаменатор: «Я тебе сейчас скажу слово, только оно не настоящее, а вымышленное, а чтобы сделать из него настоящее, ты должен убрать определенный звук. Я тебе скажу, какой. Например, если из слова "рёш" убрать "р", то получится "ёш", да? Ну, давай попробуем». **ВНИМАНИЕ:** Произносить звуки как ЗВУКИ, а не как буквы. При необходимости повторить.

Пример и проба: рёш (без "р") = ёш; шлось (без "ш") = лось

1. мон без "м"	5. брок без "р"	9. даф без "ф"
2. рум без "р"	6. твень без "в"	10. тым без "м"
3. книс без "к"	7. чакс без "к"	11. гаст без "т"
4. плёт без "п"	8. нольчь без "ль"	12. мирк без "к"

Блок 2: Тесты на Проверку Краткосрочной Памяти

Этот блок состоит из двух тестов. Тесты даются ребенку в том порядке, в каком они указаны. Порядок заданий в каждом тесте фиксированный.

Тест 1 (повторение вымышленных слов)

Экзаменатор: "Я тебе сейчас буду говорить слова, только они не нестоящие, а придуманные. Ты их постарайся повторить как можно точнее. Например, если я скажу слово "мун", ты повторишь "мун", да? Ну, давай попробуем". **ВНИМАНИЕ**: Произносить слова четко, монотонно и слитно с интервалом в одну секунду. Не повторять ни в коем случае!

1. тул	7. ру-бит	13. бар-ка-зон	19. па-но-ри-тан
2. сеп	8. пен-тел	14. бан-ди-фер	20. ти-ну-ра-зот
3. бамс	9. бал-коп	15. сти-кэ-рин	21. ам-плу-вар-синт
4. грал	10. прин-дол	16. ски-ти-культ	22. чер-кли-сте-ронк
5. хонт	11. таф-лест	17. кра-сте-рер	23. фон-тран-са-мист
6. клирт	12. сма-пент	18. гли-сте-ринк	24. клон-зер-ска-польт

Тест 2 (повторение цифр)

Экзаменатор: "Я тебе сейчас скажу несколько цифр, а ты их просто повтори. Хорошо? Например, если я скажу "три, пять", ты повторишь "три, пять", правильно? Тебе понятно, что надо делать? Ну, давай попробуем.

ВНИМАНИЕ: Произносить каждую цифру четко и монотонно, с интервалом в 1 секунду. Порядок следующий: Ряд А попытка 1, Ряд А попытка 2, Ряд Б попытка 1, Ряд Б попытка 2 и так далее. В графе "Правильно/Нет" поставить плюс, если ответ правильный и минус, если ответ неправильный.

Пример и проба: 35 (три, пять); 256 (два, пять шесть)

Ряд	Попытка 1	Правильно/Нет	Попытка 2	Павильно/Нет
А	43		16	
Б	792		847	
В	5941		7253	
Г	93872		75396	
Д	152649		216748	
Ж	3745261		4925316	
З	82973546		69174253	
И	246937185		371625948	

Блок 2: Тесты на Проверку Автоматического Доступа

Блок состоит из двух тестов. Тесты даются ребенку в том порядке, в каком они указаны.

Тест 1 (картинки)

Покажите ребенку картинки, используемые в этом тесте, и спросите, "Как называются эти предметы?" Если ребенок правильно назвал все предметы, положите перед ним Приложение 1 с рядами картинок так, чтобы он видел только первые два ряда и скажите: "А теперь назови эти картинки подряд как можно быстрее." Убедившись, что ребенок понимает, что от него требуется, откройте все картинки и скажите: "А теперь назови все эти картинки с начала и как можно быстрее." Засеките время на секундомере. После окончания задания, проставьте время внизу в строке «Время».

ВНИМАНИЕ: Если предложенные ребенком названия отличаются от названий в этом тесте (например, ребенок говорит "зонтик" вместо "зонт"), исправьте ребенка. Называть картинки на листе следует горизонтально, в ряд. Для этого теста используется Приложение 1.

Правильные названия картинок: стул, кость, торт, мяч, зонт

Время:_____

Тест 2 (цифры)

Покажите ребенку цифры, используемые в этом тесте, и попросите их назвать. Если ребенок правильно их назвал, положите перед ним Приложение 2 с рядами цифр так, чтобы он видел только первые два ряда и скажите: "А теперь назови эти цифры подряд как можно быстрее." Убедившись, что ребенок понимает, что от него требуется, откройте все цифры и скажите: "А теперь назови все эти цифры с начала и как можно быстрее." Засеките время на секундомере. После окончания задания, проставьте время внизу в строке «Время».

ВНИМАНИЕ: Если предложенные ребенком названия отличаются от названий в этом тесте (например, ребенок говорит "двойка" вместо "два"), исправьте ребенка. Называть цифры на листе следует горизонтально, в ряд. Для этого теста используется Приложение 2.

Правильные названия цифр: два, три, пять, шесть, семь

Время:_____

Блок 4: Чтение Слов

Экзаменатор: «Прочитай мне, пожалйуста, эти слова. Читай их в строчку, а не в столбик, и так быстро, как только можешь, но без ошибок. Номера не произноси.» Для этого теста используются листы в Приложении 3. Время засекается для каждого листа отдельно. Порядок листов и слов фиксированный и одинаковый для всех. Дайте ребенку сначала один лист, засеките время, а после окончания чтения, проставьте это время. Потом дайте ему второй лист и снова засеките время. Следует записать неправильно произнесенные слова так, как ребенок их произнес. По возможности ответы записать на магнитофон.

Блок 5: Чтение Вымышленных Слов

Экзаменатор: «Давай мы с тобой еще почитаем. Вот здесь написаны слова, но только они не настоящие, а придуманные. Попробуй их прочитать. Постарайся читать быстро и аккуратно, без ошибок. Читай в строчку, по номерам. Номера не произноси». Для этого теста используются листы в Приложении 4. Время засекается для каждого листа отдельно. Порядок листов и слов фиксированный и одинаковый для всех. Дайте ребенку сначала один лист, засеките время, а после окончания чтения, проставьте это время. Потом дайте ему второй лист и снова засеките время. Следует записать неправильно произнесенные слова так, как ребенок их произнес. По возможности ответы записать на магнитофон.

Блок 6: Понимание Прочитанного на Слух

Экзаменатор: «Я тебе сейчас прочитаю несколько рассказов. Ты их послушай, а потом я тебе задам по ним несколько вопросов. Постарайся на них оветить. Хорошо?» **ВНИМАНИЕ:** Названия НЕ читать. После каждого рассказа задавать вопросы и коротко записывать ответы. Если ответ непонятен, задать уточняющие вопросы. Читать каждый рассказ один раз.

Рассказ 1: Белка

Мы с папой гуляли в парке. Вдруг на дорожку выскочил пушистый зверек. Это была ручная белка. Я бросил ей орех. Она схватила его и быстро зарыла под сосной. Так белочка делала свои запасы.

Вопросы

1. Где произошел этот случай?
2. Зачем белка зарыла орех?

Рассказ 2: Пингвины

Кто такие пингвины? Пингвины – это птицы. Но птицы особенные. Они не умеют летать. Пингвины хорошо плавают. Пищу они добывают в воде, а большую часть суток проводят на суше.

Вопросы

1. Кто такие пингвины?
2. В чем особенность этих птиц?
3. Где они добывают себе пищу?

Рассказ 3: Устная речь

Люди стали говорить очень давно. Им нужно было добывать пищу, строить жилища. Сообща это делать проще. И люди звали друг друга на помощь. Так появилась устная речь.

Вопросы

1. О чем говорится в этом рассказе?
2. Как появилась устная речь?
3. Для чего люди звали друг друга на помощь?

Рассказ 4: Три товарища

Витя потерял завтрак. На большой перемене все ребята завтракали, а Витя стоял в сторонке.

- Почему ты не ешь? – спросил его Коля.
- Я завтрак потерял.
- Плохо, - сказал Коля, откусывая большой кусок белого хлеба.
- А где ты его потерял? – спросил Миша.
- Не знаю, – тихо сказал Витя и отвернулся.
- Ты, наверное, в кармане нёс, а надо в сумку класть, - сказал Миша.

А Володя ничего не спросил. Он подошел к Вите, разломил пополам кусок хлеба с маслом и протянул товарищу:

- Бери, ешь!

Вопросы

1. Где произошел этот случай?
2. Почему Витя стоял в сторонке и не завтракал?
3. Сколько товарищей подошли к нему?
4. Помог ли кто-нибудь из них Вите?

Рассказ 5: Письмо

Серёжа получил письмо от друга. Он живёт в Москве. Андрей пишет о своём классе. У него хорошие друзья. Недавно у него был день рождения, и ему подарили коллекцию марок.

Вопросы

1. Кто получил письмо от друга?
2. Где живёт Андрей?
3. О чём он пишет в письме?
4. Что ему подарили на день рождения?

Блок 7: Матрицы (Невербальный Интеллект)

Для этого теста используется Субтест Матрицы Теста KBIT-2 (Kaufman, A., & Kaufman, N. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Circle Pines, MN: AGS.).

Общие правила

1. Попросите ребенка показать на картинку, которую он считает самой подходящей в каждом вопросе. Никакие подсказки не допускаются.
2. Тест прекращается после четырех неправильных ответов подряд.
3. Если ребенок слишком долго думает над вопросом, скажите: "Просто выбери любую картинку" и идите дальше.
4. Ответы ребенка (латинские буквы указанных картинок) указать в Ответном Листе в Приложении 5.

Правила для отдельных вопросов

Sample A: «Эта картинка (покажите на машину) подходит к одной из этих (покажите на весь ряд). Покажи, к какой из них она больше всего подходит». Если ребенок правильно ответил, переходите к вопросу 1, если нет, объясните следующее: "Это машина. Она подходит к грузовику (показывая). У обоих есть колеса и люди на них ездят. Ну, давай еще раз». Повторите этот же вопрос.

Вопрос 1: «Эта картинка (показывая на туфлю) к какой больше всего подходит?» Если ответ неправилен, объясните, почему эта картинка подходит к картинке ноги—люди носят туфли на ногах. Повторите этот вопрос.

Вопрос 2: «Эта картинка (показывая на монеты) к какой больше всего подходит?» При неправильном ответе, объясните, что монеты подходят к банкнотам—и те и другие—деньги. Повторите этот вопрос.

Вопросы 3-9: «Эта картинка (показывая на верхнюю) подходит больше всего к какой из этих (показывая на ряд)»? Запишите ответ ребенка. **Никаких подсказок и объяснений здесь не давать.**

Sample B: «Эта картинка (показывая на морковь) подходит к этой (кролик), так же как и эта (кость) подходит к какой из этих (нижний ряд)»? При неправильном ответе объясните: Кролик ест морковь, а собака—кость. Повторите вопрос.

Вопрос 10: «Эта картинка (машина) подходит к этой (дорога) так же как и эта (яхта) подходит к какой из этих (нижний ряд)»? При неправильном ответе, объясните, что машина ездит по дороге, а яхта—по воде. Повторите этот вопрос.

Вопрос 11: «Эта картинка (кролик) подходит к этой (кроличьи уши), так же как и эта (слон) подходит к какой из этих (нижний ряд)»? При неправильном ответе, объясните, что уши кролика выглядят так (показывая на картинку), а уши слона—так (показывая на картинку). Повторите вопрос.

Вопросы 12-22: Покажите на знак вопроса и спросите "Какая картинка сюда подходит?" Никаких подсказок и объяснений не давать!

Sample C: «Все эти кусочки (покажите на нижний ряд) имеют одинаковую форму. Какой из них подходит сюда»? (Покажите на знак вопроса.) При неправильном ответе объясните: В этом ряду (верхний) мячи такие, а в этом (нижний)—такие. Поэтому в этом ряду все мячи должны быть такие (покажите на картинку футбольного мяча). Повторите этот вопрос.

Вопрос 23: Покажите на знак вопроса и спросите: "Какая картинка сюда подходит?" При неправильном ответе объясните: "Весь этот ряд (левый)—велосипеды, а этот (средний)—грузовики. Поэтому этот ряд должен быть воздушные шары" (показывая на правильный ответ). Повторите этот вопрос.

Вопрос 24: Покажите на знак вопроса и спросите: "Какая картинка сюда подходит?" При неправильном ответе объясните: "В верхнем ряду, вертикальные линии широко расставлены, а фигуры—треугольник, квадрат и круг. В среднем ряду вертикальные линии ближе друг к другу, а фигуры такие же. Поэтому в этом ряду (нижний) вертикальные линии должны быть еще ближе друг к другу, а в центре должен быть квадрат. Поэтому правильный ответ—этот" (покажите на картинку F. Повторите вопрос.

Вопросы 25-46: Покажите на знак вопроса и спросите: "Какая картинка сюда подходит?" Никаких объяснений не давать!

Блок 8: Словарный Тест

В этом блоке используется Словарный Субтест теста Векслера, русский вариант (Филимоненко, Ю., и Тимофеев, В. (2001). *Тест Д. Векслера: Диагностика структуры интеллекта (детский вариант). Методическое руководство*. Санкт-Петербург: Иматон, стр. 30-43.).

Блок 9: Понимание Прочитанного

Экзаменатор: «Давай мы с тобой немножко почитаем. Вот здесь есть короткие рассказы, ты их читай вслух, а я потом тебе задам по каждому несколько вопросов. Хорошо?» Рассказы в Приложении 6. Ребенок читает **вслух**. Во время чтения необходимо сразу поправлять неправильно прочитанные слова. После начала чтения, засечь время; после окончания, проставить его в строке «время». После чтения каждого рассказа, задавать вопросы (внизу). По возможности ответы записать на магнитофон.

Рассказ 1: Пушок
Время:_____

1. О ком говорится в этом рассказе?
2. Почему ежа прозвали Пушком?

Рассказ 2: В детском саду
Время:_____

1. Что делали ребята в детском саду?
2. Почему Дима не играл?
3. Почему ему было скучно?

Рассказ 3: Правда всего дороже
Время:_____

1. Что сделал мальчик?
2. Почему он затрясся от страха на вопрос отца?

3. Почему отец сказал спасибо?

Рассказ: Барсук
Время:_____

5. Что заставило барсука обратиться за помощью к человеку?
6. Как голова барсука оказалась в банке?
7. Куда прибежал зверёк?
8. Как лесник помог бедняге?

Рассказ 5: Колибри и слон
Время:_____

1. Какая птичка самая маленькая?
2. Какого она размера?
3. Какой высоты достигает слон?
4. Где живут крошечеая птичка и огромный слон?

Приложение 1

Приложение 2

2	3	5	6	7
5	2	3	7	6
5	6	7	3	2
7	3	6	5	2
3	2	5	6	7
6	7	2	5	3
3	5	7	2	6
5	3	6	2	7
2	3	6	7	5
7	5	3	2	6

Приложение 3

Лист 1

1. лес	2. пень
3. крот	4. волк
5. тигр	6. весна
7. доска	8. идёт
9. якорь	10. друзья
11. серьёзно	12. поздравил
13. сплоховал	14. множество
15. подснежник	16. неизвестно
17. спрашивает	18. прожужжала
19. заблудилась	20. подглядывать

Лист 2

1. бой	2. сон	3. дуть	4. мчал
5. корм	6. шуба	7. дело	8. ссора
9. голос	10. чувство		

Лист 3

1. разбойник	2. спросонья	3. надувной	4. помчался
5. кормушка	6. полушубок	7. безделушка	8. перессорить
9. голосистый	10. почувствовал		

Лист 4

1. коридор	2. пластилин
3. превратить	4. отшвырнул
5. встревожить	6. пожалуйста
7. отечество	8. разъярился
9. действительно	10. всполошились

Приложение 4

Лист 1

1. сой	2. бон
3. корс	4. снем
5. стур	6. дроль
7. башай	8. калей
9. кисан	10. чигра
11. леньтю	12. дышлан
13. безарё	14. изкабуш
15. прунажи	16. кадашран
17. путыйшис	18. буникдиль
19. пенерифул	20. финоризер
21. контрапонист	22. перплистеронк
23. эмплифорвент	24. блонтерстапинг

Лист 2

1. лас	2. наж
3. кень	4. дыль
5. тугр	6. сталб
7. авощ	8. вьёга
9. завыд	10. юскра
11. зувисть	12. встроча
13. магазон	14. конфата
15. коруткий	16. дровосяк
17. спрытался	18. подпрагнул
19. волныние	20. удивалась
21. эскалутор	22. раскалянный
23. предупредал	24. тропуческий

Приложение 5

Поставьте в графе "Ответ" букву, соответствующую картинке, на которую указал ребенок. Прекратить тест после 4 неправильных ответов **подряд.**

Вопрос	**Правильный ответ**	**Ответ ребенка**
Sample A	A	
1	C	
2	B	
3	D	
4	E	
5	D	
6	E	
7	A	
8	B	
9	D	
Sample B	C	
10	B	
11	D	
12	F	
13	B	
14	E	
15	C	
16	C	
17	B	
18	A	
19	D	
20	A	
21	E	
22	D	
Sample C	C	
23	B	
24	F	
25	E	
26	A	
27	C	
28	E	
29	F	
30	F	
31	E	
32	A	
33	F	
34	C	
35	E	
36	E	
37	D	
38	A	
39	F	
40	B	
41	C	
42	F	
43	C	
44	A	
45	B	
46	A	

Приложение 6

Рассказ 1

В доме у нас жил ёжик. Он был ручной. Когда его гладили, он прижимал к спине колючки и делался совсем мягким. За это мы его прозвали Пушком.

Рассказ 2

В детском саду было много игрушек. Ребята играли все вместе. Им было весело. Только Дима не играл. Он собрал около себя целую кучу игрушек и охранял их от детей. “Моё! Моё!” – кричал он, закрывая игрушки руками. Ему было скучно.

Рассказ 3

Мальчик играл и разбил нечаянно дорогую чашку. Никто не видал. Отец пришел и спросил: “Кто разбил?” Мальчик

затрясся от страха и сказал: “Я”. Отец сказал: “Спасибо, что правду сказал. Правда всего дороже”.

Рассказ 4

Однажды к избе лесника из чащи прибежал барсук. На голове у него стеклянная банка надета, как скафандр. Смотрит барсук сквозь банку, а снять ее не может. Что-то вкусное было в банке, вот зверёк и попался. За помощью он прибежал к человеку. Лесник аккуратно разбил банку и освободил беднягу. Счастливый барсук быстро скрылся в ельнике.

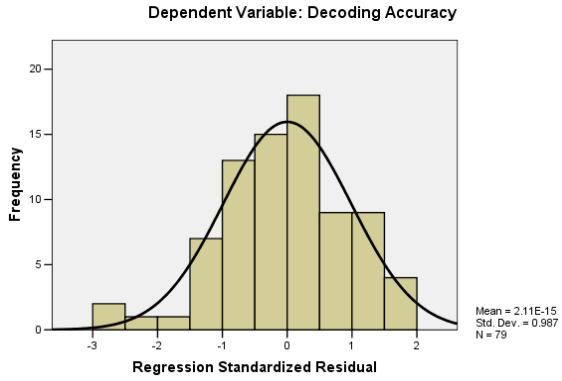
Рассказ 5

Самая маленькая птичка – колибри. Она размером с бабочку. Самое большое животное на Земле – слон. Он достигает высоты до трёх метров. Крошечная птичка и огромный слон живут в тёплых странах.

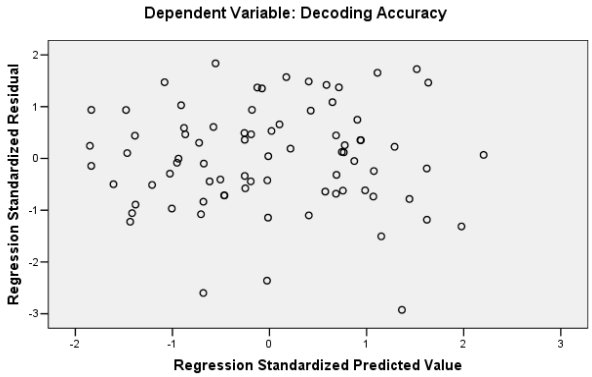
APPENDIX Q

RESIDUALS SCATTERPLOTS FOR DECODING ACCURACY

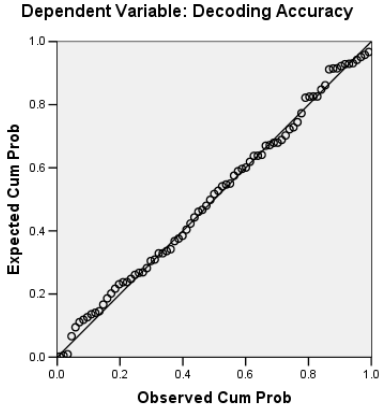
Histogram



Scatterplot



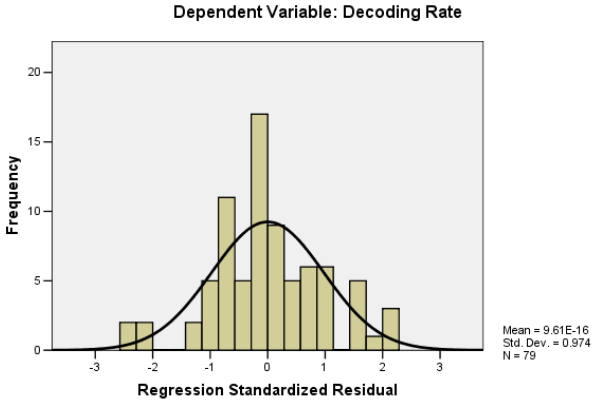
Normal P-P Plot of Regression Standardized Residual



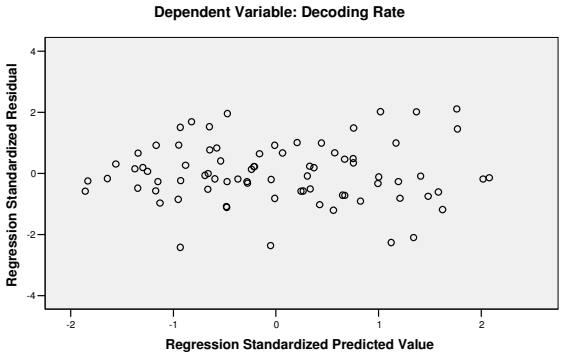
APPENDIX R

RESIDUALS SCATTERPLOTS FOR DECODING RATE

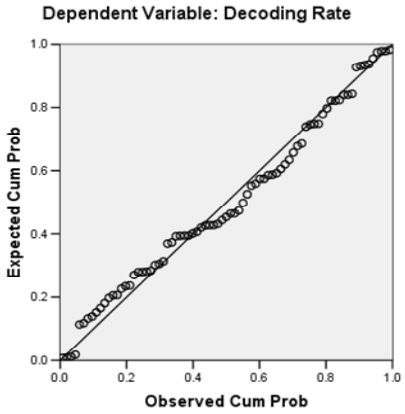
Histogram



Scatterplot



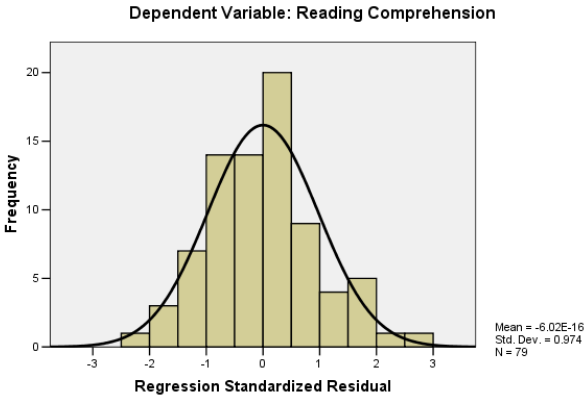
Normal P-P Plot of Regression Standardized Residual



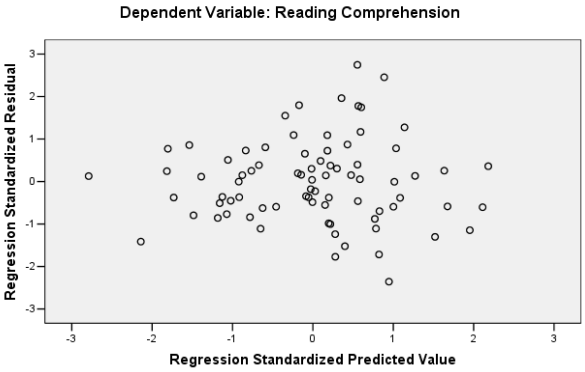
APPENDIX S

RESIDUALS SCATTERPLOTS FOR READING COMPREHENSION

Histogram



Scatterplot



Normal P-P Plot of Regression Standardized Residual