G Model
EARCHI-586; No. of Pages 16
ARTICLE IN PRESS
Early Childhood Research Quarterly xxx (2012) xxx–xxx
Contents lists available at SciVerse ScienceDirect
Early Childhood Research Quarterly

Children learning to read later catch up to children reading earlier

Sebastian P. Suggate[a,*], Elizabeth A. Schaughency[b], Elaine Reese[b]

[a] University of Regensburg, Germany
[b] University of Otago, New Zealand

ARTICLE INFO

Article history:
Received 30 March 2011
Received in revised form 5 April 2012
Accepted 22 April 2012

Keywords:
Reading age
Reading achievement
Phonemic awareness
Reading comprehension
Early reading

ABSTRACT

Two studies from English-speaking samples investigated the methodologically difficult question of whether the later reading achievement of children learning to read earlier or later differs. Children ($n = 287$) from predominantly state-funded schools were selected and they differed in whether the reading instruction age (RIA) was either five or seven years. Study 1 covered the first six years of school following three cohorts across a two-year design. Analyses accounted for receptive vocabulary, reported parental income and education, school-community affluence, classroom instruction, home literacy environment, reading self-concept, and age. The earlier RIA group had initially superior letter naming, non-word, word, and passage reading but this difference in reading skill disappeared by age 11. In Study 2, the decoding, fluency, and reading comprehension performance of 83 additional middle school-age children was compared. The two groups exhibited similar reading fluency, but the later RIA had generally greater reading comprehension. Given that the design was non-experimental, we urge further research to better understand developmental patterns and influences arising from different RIAs.
© 2012 Elsevier Inc. All rights reserved.

1. Introduction

The age at which children enter school and learn to read is intuitively an important factor in later reading achievement. Formal reading instruction typically begins comparatively early in the English-speaking world, around age five, with there being a recent trend for other countries—and early education institutions within countries—to lower the age when children begin. In the US, for example, the call for early intervention to address risk for reading failure was repeatedly made at hearings for the National Reading Panel (2000), with the publication of this report credited as contributing to initiatives such as Reading First, focusing on reading instruction in the early grades so that children become competent readers by the end of third grade, and Early Reading First, intended to promote reading readiness skills among preschool-age children (Shanahan & Lonigan, 2010). Similarly, in England, literacy is included among the early learning goals for the preschool years (Quick, Lambley, Newcombe, & Aubrey, 2002), with many preschool (i.e., reception) classrooms providing a Literacy Hour by the end of the school year (Quick et al., 2002). This practice is not without controversy, however: An independent review of English elementary education criticized the early start to formal schooling and recommended re-examining school starting age (Cambridge Primary Review, 2009). Moreover, a report on reading instruction for the House of Commons concluded that current evidence does not provide a clear answer to the question of what is the best age to begin formal teaching (House of Commons Education and Skills Committee, 2005).

1.1. Previous research

Contemporary researchers have typically considered the question of whether preschool or kindergarten children are too young to learn to read (e.g., Cunningham & Carroll, 2011a; Ehri, Nunes, Stahl, & Willows, 2001). The answer to this question appears to be "no" for many children. Evidence from meta-analysis suggests that preschool and kindergarten children, including those at-risk, can be taught decoding skills (e.g., Bus & van Ijzendoorn, 1999; Ehri, Nunes, Stahl et al., 2001; Suggate, 2010). Moreover, single studies (Feitelson, Tehori, & Levinberg-Green, 1982) and comparative international studies (Seymour, Aro, Erskine, & COST Action Network, 2003) suggest that younger children can learn to read from the age of four.

However, in this paper we turn to a somewhat different question: Is there a long-term advantage in reading for having learned to read earlier? Given that the reading skills gained in elementary school predict subsequent educational and occupational pathways for children (Savolainen, Ahonen, Aro, Tolvanen, & Holopainen, 2008), coupled with the poignant reality that many children fail

* Corresponding author at: Lehrstuhl für Schulpädagogik, University of Regensburg, Universitätsstraße 31, 93040 Regensburg, Germany. Tel.: +49 941 943 1750; fax: +49 941 943 1993.
E-mail addresses: sebastian.suggate@ur.de, sebastian.suggate@gmail.com (S.P. Suggate).

0885-2006/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ecresq.2012.04.004

Please cite this article in press as: Suggate, S. P., et al. Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly* (2012), http://dx.doi.org/10.1016/j.ecresq.2012.04.004

to master these skills (Adams, 1990; National Early Literacy Panel (NELP, 2008) ; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998), this question has educational significance. Also, given that reading involves the weaving together of pre-reading, decoding, reading, and language-related skills (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Whitehurst & Lonigan, 2001), skills that may be differentially influenced by formal and informal learning experiences (Cunningham & Carroll, 2011a), examination of the effect of reading instruction age (RIA) potentially enhances our understanding of reading development.

We recognize that skilled reading typically comprises the development of pre-reading, decoding, reading, and reading comprehension skills, each of which is influenced by a variety of preschool and school experiences (Adams, 1990; Biemiller, 2006; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Neuman, 2006; Storch & Whitehurst, 2002; Whitehurst & Lonigan, 2001). Because it is widely recognized that reading comprises two main strands of text-related and language-related skills (Gough & Tunmer, 1986), discussions of the sub-components of pre-reading, reading, and reading comprehension skills typically include both language- and text-related skills. Given that we seek to investigate whether the earlier introduction of formal reading instruction and concordant earlier mastery of reading leads to later advantage in reading, we are left with a definitional dilemma: On the one hand, in framing our research question, we do not believe it conceivable to test whether pre-literacy skills should be either taught in their entirety or removed altogether. There is a solid research base supporting the importance of oral language and phonemic awareness skills in reading development (e.g., NELP, 2008); yet on the other hand we seek to study the role of formal reading instruction.

Accordingly, we adopt a similar distinction to the Simple View of Reading (SVR; Gough & Tunmer, 1986) and apply this to pre-literacy and early reading development. Thus, here we define decoding-related skills as those that require text, including alphabetic awareness, decoding skills, and reading fluency. We call our second strand of skills language comprehension skills, consistent with the SVR (Gough & Tunmer, 1986), including phonemic awareness and phonology, vocabulary, syntax, semantics, pragmatics, and thinking skills used in understanding written and spoken language. Although this definition is not perfect because many decoding-related skills also require language, it captures the essence of the distinction between early instruction of decoding-related skills while acknowledging that oral language and phonemic awareness belong to pre-literacy and reading development. Finally, we define skilled reading as being able to read accurately and fluently, and to understand what is read.

Herewith, we also adopt the view that many important pre-literacy experiences constitute, in some respects, the beginning of learning that leads to competent reading skills (Whitehurst & Lonigan, 2001). However, the question we ask here is whether the early development of decoding-related skills per se is important in later skilled reading? We define early reading instruction as that intended to lead to the development of decoding-related skills, be they more focused on teaching the explicit relations between graphemes and phonemes or between deciphering graphemes with the help of contextual cues. In doing so, we do not relegate the role of language in reading (Dickinson et al., 2010), but rather seek to isolate the role of early reading instruction in the later development of skilled reading.

Knowledge of decoding-related skills, including letter knowledge, letter-sound correspondences, and word-reading, has proven to be a robust predictor of later reading in correlational studies (Snow et al., 1998). For example, preschool or kindergarten non-word decoding correlates highly with later reading ($r = .72$, Lonigan, Schatschneider, Westberg, & NELP, 2008). Importantly, because a variety of third factors are associated with, and may well determine and drive, both early and later reading skill—such as instruction (National Reading Panel, 2000), home environment (Niklas & Schneider, 2010), genetic (Olson & Gayan, 2001), and language factors (Dickinson et al., 2003)—correlations between early decoding-related skills and later achievement do not fully answer the question of how early versus later readers fare long-term.

A second, important line of research arises from controlled reading intervention studies. Many intervention studies have shown that early decoding-related skills can be boosted in the short-term (Bus & van Ijzendoorn, 1999; Ehri, Nunes, Stahl et al., 2001; Ehri, Nunes, Willows et al., 2001; Suggate, 2010). However, seldom do studies include long-term data beyond 12 to 18 months post-intervention, preventing investigation of whether the early advantage in skilled reading later washes out. Moreover, such intervention studies usually compare children receiving one form of decoding-related instruction with another form. Thus, even the control children in intervention studies in kindergarten or grade 1 learned early decoding-related skills or received early instruction. Clearly, to understand whether early reading instruction leads to long-term advantage in reading, two groups differing in when they receive reading instruction need to be compared over a long period of time.

Few studies directly compare the long-term effects of having an earlier versus later reading instruction age (RIA). Findings from two international studies suggest that children with a later school entry age—which can be expected to correlate highly with RIA—achieve similarly in long-term reading achievement (Elley, 1992; Suggate, 2009). However, by virtue of being cross-sectional, these studies do not account for development at the beginning of primary school and conflate international language differences with RIA. Importantly, the countries with the earlier RIAs also speak English, which has a uniquely complex orthography and spelling-to-sound structure (Seymour et al., 2003). International studies do not rule out the possibility that an earlier RIA is needed to offset the added difficulty in learning to read English.

We consider that a stronger test of whether there is a long-term relative advantage in reading achievement from earlier reading instruction would involve comparisons of children beginning formal reading instruction at markedly different ages with a follow-up assessment occurring at a chronologically equivalent age. Therewith children beginning at five, the typical age of beginning reading instruction in English-speaking countries, could be contrasted with children beginning at age six or seven, which more closely approximates the typical age of beginning reading instruction in other highly literate countries, such as those in Scandinavia. Moreover, given English's unique orthographical features, a strong test would involve English-speaking samples and, given the difficulty of random assignment, control for key variables, and capture learning longitudinally.

The ethics of conducting this methodologically ideal study currently precludes random assignment because evidence from less invasive non-experimental research would first need to be compelling. In a cross-sectional study in the United States, children randomly selected from a waiting list to either enter or have entry declined to a Montessori school were compared (Lillard & Else-Quest, 2006). In Montessori schools, children often end up learning to read sometime soon after age three, perhaps because of the individual support offered them (Edwards, 2007). Accordingly, the Montessori children had an advantage in reading over the public-schooled children at age five but not at age 12. An older study (i.e., Durkin, 1974–1975) also investigated reading longitudinally but had a small sample size, did not have random assignment, focused on IQ as the main potential confound, and covered only two years of the primary school period. Interestingly, the findings of this study were consistent with those of Lillard and Else-Quest (2006) and the trends from international studies (Elley, 1992; Suggate, 2009): After

Please cite this article in press as: Suggate, S. P., et al. Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly* (2012), http://dx.doi.org/10.1016/j.ecresq.2012.04.004

two years of primary school, the children whose parents opted for entry into the early school program had lost their earlier advantage in reading skill.

A more recent study conducted in the UK capitalized on schooling differences between Steiner and standard state curricula (described more fully below) to investigate whether the younger state children would acquire beginning decoding-related skills more slowly than the older Steiner-educated children (Cunningham & Carroll, 2011a). Decoding-related development was followed across the first year of school for both the state-educated ($n = 31$) children aged four years and nine months and the Steiner-educated children ($n = 30$) aged seven years and 10 months. Follow-up data were collected after two years of school on a smaller sample ($n = 19$ for Steiner educated and $n = 25$ for state educated). Key findings were: (a) across the initial two years of decoding-related instruction, there were no systematic differences in reading performance, (b) the state-educated children had an advantage in spelling, and (c) the state-educated children received more reading and phonics instruction. Moreover, Cunningham and Carroll's (2011a) study does not explore the question of whether the earlier mastery of decoding-related skill by the state-educated sample would lead to a *later* advantage in skilled reading on a chronologically matched sample for two main reasons. First and foremost, the study was not conducted sufficiently long-term to allow for age-matched comparisons. Second, importantly, given the association of socioeconomic factors (Chiu & McBride-Chang, 2006) and parental education (Boyce & Snow, 2009) to reading achievement, Steiner education in the UK is privately funded and there appeared to be marked differences in parental education between the two samples in Cunningham and Carroll's study. Therefore, given differences in parental education and financial constraints on school attendance, the extent to which factors related to the cultural capital within families contributed to reading outcomes is uncertain.

1.2. Background to the current studies

To address the question of whether early RIA is associated with later advantages in reading achievement, we present the results of two studies, one employing three pairs of longitudinal samples and the other cross-sectional, conducted in New Zealand (NZ), spanning the first six years of school, for students with an RIA of five versus seven years. By spanning the first six years of school, the design captured a diverse period of reading acquisition, enabling consideration of the antecedents of reading, fluency development, and later reading comprehension (Leppänen, Aunola, Niemi, & Nurmi, 2008). Moreover, NZ presents an opportune context for investigating the role of the RIA because the same two educational options as in Cunningham and Carroll (2011a, 2011b) studies are available, which similarly differ in RIA, but with one crucial difference: Government funding allows parents to enroll children in either type of school, often without significant differences in parental financial contributions.

1.2.1. Reading instruction in state schools

Children who attend state schools in NZ begin school on their fifth birthday (unless this date occurs during a holiday) and receive formal reading instruction from that day. As described in Ministry of Education teacher resource materials (Ministry of Education, 2007), literacy-related activities should pervade the school day in the first year of school. The typical reading session is described as involving shared and guided reading in a language-rich text (Ministry of Education, 2007). Context-based word identification strategies have influenced approaches to reading instruction in NZ state schools (Clay, 1998; Smith & Elley, 1997; see Greaney, 2002, for discussion), although Ministry of Education documents also discuss recommendations for greater attention to word-level decoding-related skills as an aspect of reading acquisition (Ministry of Education, 1999, 2003b, 2010) along with development of a sight reading vocabulary (Ministry of Education, 2003a, 2010).

Although the approach in NZ classrooms may not always include explicit instruction in phonemic awareness and phonics as recommended in recent syntheses such as that from the National Reading Panel (2000), scores on leveled readers indicate that, by age six, most children can read connected text consisting of about 300 words with one storyline or topic with an accuracy of over 90 percent (McNaughton, Phillips, & MacDonald, 2000). This normative standard of reading at level 12 (Green) in the *Ready to Read* series of leveled readers has been set as the benchmark for achievement in national achievement standards being implemented in 2010 (Ministry of Education, 2009), with the expectation that children who are achieving below the 15th percentile should be provided with additional reading support, such as through one-to-one remedial programs (Ng, 2006). Our previous research in NZ suggests that many children entering school at age five are just beginning to develop alphabetic skills (e.g., correctly naming, on average, approximately 12 letters in one minute, Suggate, Schaughency, & Reese, 2011); therefore we consider it justified to assume a formal RIA of five years for state-curriculum children, while acknowledging that some children may have acquired literacy, and even reading, skills in preschool or at home. Moreover, given that, by age six, many children meet expectations for text reading (e.g., using punctuation marks to produce smooth phrasing in oral reading), we consider also that, in addition to receiving early reading instruction, many NZ children accurately read connected text with expression.

1.2.2. Reading instruction in Steiner education

Conversely, children in the Steiner curriculum enter school in the year of their seventh birthday, typically spending the prior two years in Steiner kindergartens instead of state schools. In these kindergarten years, *written* language—be it in the form of letters, words or stories—is excluded from the kindergartens to encourage the development of *oral* language. Instead, children's activities center on play, activities (e.g., painting, drawing, cooking), singing or oral story telling from the kindergarten teacher (Edwards, 2007). Therefore, reading instruction does not occur in Steiner kindergartens. However, we recognize that the kinds of language activities taking place may well develop a strong foundation for later skilled reading (Dockrell, Stuart, & King, 2010), even though they do not involve text.

Unlike in state schools in NZ where most children enter school when they are precisely five years old, Steiner children enter at the beginning of the academic year when they turn seven. Accordingly, they have an average chronological age of 6.5 years, hence the difference in RIA with state school children is around 18 months. Reading instruction begins in the first year of school through the artistic introduction of the letters and their corresponding sounds through story, one per week (Steiner, 1924), which lasts until the middle of grade 1. Once children have acquired a number of consonants and vowels, they begin exploring how words comprise smaller phonemic/graphemic components, with a focus on segmenting as opposed to blending (Burnett, 2007). Children also gain experience with the letters through writing and forming the letters they know into words (Burnett). Towards the end of grade 1 and in grade 2, children begin learning to read connected text. Reading activities early in the Steiner-curriculum have been described as following an analytic phonics approach, as this tends to focus on word recognition through learning of initial sound letter correspondences and sight word reading (Cunningham & Carroll, 2011a). This less synthetic and systematic approach to phonics in the Steiner schools, coupled with sight recognition training and use of oral language to identify words, gives rise to more similarity between the NZ state and Steiner curriculum than is the case in the UK.

Moreover, similarities between the NZ state and Steiner curricula exist after the initial acquisition of decoding-related skills.

1.3. The current studies

Capitalizing on these RIA differences, the first study follows reading development across the first six years of school, which corresponds to the typical length of elementary school in NZ. The second study extends the first by focusing (a) on an older Year-7 state school (i.e., beginning middle school) sample that is closer in age to the Year 6 end-point Steiner sample in the first study, (b) on reading comprehension, and (c) at an age when children are expected to be reading for meaning (Chall, 1987; Ministry of Education, 2009).

To begin to explore the extent to which curricula in state- and Steiner-curricula RIAs covary, we include a measure of classroom instruction. We also thought it important to account for a host of family and child characteristics, to evaluate whether the Steiner and state curricula children were comparable on variables suggested by previous research to be related to reading achievement.

Finally, a third feature of the first of the current studies is that we build upon the conclusions and recommendations for future research made by members of the NELP (Shanahan & Lonigan, 2010). Recognizing that young children change over development, we followed reading trajectory development at regular intervals, investigating conventional pre-literacy, decoding, and reading skills (i.e., decoding, oral reading accuracy and fluency, and reading comprehension) identified in the National Reading Panel (2000) and NELP (2008) reports. By assessing development of decoding-related skills, we are able to compare groups as a function of age and exposure to schooling and different educational experiences, thus extending the work of Cunningham and Carroll (2011a, 2011b).

2. Study 1

Study 1 longitudinally follows three pairs of elementary samples to investigate the long-term development of reading skill as a function of RIA. We further investigated for differential development of specific reading skills, expecting to see initial greater decoding-related skills for the early RIA sample, given previous research (Cunningham & Carroll, 2011a; Suggate et al., 2011). Moreover, we included both comprehension and reading outcome measures for the older children, to ensure that reading skill development was not confined to more constrained measures (Paris, 2005), which could potentially lead to ceiling effects, masking group differences.

Participants in Study 1 were children in their first (junior cohort), third (middle cohort), or fifth (senior cohort) year of primary school at the beginning of the study and followed across two years, thus providing the opportunity to observe reading performance across the period of elementary education in New Zealand. Crucially, the key independent variable of RIA was determined by whether children were recruited from state schools (RIA of five years) versus Steiner schools (RIA of seven years). Table 1 presents the overall design of the study, including the cohorts and measurement schedule.

2.1. Method

2.1.1. Participants

Participants were recruited from three government-supported Steiner curriculum ($n = 111$) and three state curriculum schools ($n = 234$). Technically, two Steiner schools were "state integrated," receiving full governmental funding but allowed to retain their special character. The third school was an "independent school" which also received some governmental funding, but at a lower rate (not exceeding 30%). This independent Steiner school ($n = 21$) was included in this research to bolster sample size and to improve representation of lower SES families, as this school drew its children from a community with a low to moderate SES. Government-reported indices suggested no state/Steiner differences in estimates of the economic affluence of school communities (see discussion of school decile, below), increasing the likelihood that the samples were matched on SES.

During the course of the study, 14.96 percent of state and 9.91 percent of Steiner children left the schools, with the difference between school types not being significant according to a Fisher's exact test (two-tailed, $p = .24$). One of these children left the study near its completion, such that these data were still able to be included in the analyses. The reasons for leaving the schools appeared to be due to a residential address change, according to teacher reports. A further 13 children's data were eliminated because: (a) either there was too much missing data, defined as three or more missing data points, or (b) the children were given the wrong test materials, usually due to their having been in a lower or higher academic year than that tested, with this placement changing in the course of the study (due to staggered school entry according to birthdates, some children enter Year 1 directly, whereas others enter the year previously denoted as Year 0). This left 190 state and 97 Steiner pupils, 13 of whom first entered state schooling and, therefore, had an RIA of five even though they later attended steiner schooling. The reasons why 13 children in our Steiner school samples had changed from a state school are not known in every instance, but include dissatisfaction with the former school, and waiting for an available place. Accordingly, two variables ensue, either (a) curriculum, with children either in the Steiner versus state curriculum during the study, or (b) RIA, with children either beginning school earlier in the state school—regardless of whether they transferred into a Steiner school—or later and entering formal schooling for the first time in the Steiner curriculum around age seven.

The early RIA children began school at 60.09 months ($SD = .44$) and the later RIA, on average, 19 months later ($M = 78.64$, $SD = 4.02$). Forty-one of the 80 Year 1 children attended school in 2006, since their fifth birthday was prior to 2007 (in NZ, school entry typically occurs on the day of the fifth birthday), such that they entered school at the normal age of 5 in 2006, but entered the study when a little older at the beginning of 2007. Participation rates were estimated from Ministry of Education and school roll data as 83.72 percent for children in the Steiner-curriculum and 59.77 percent in the state curriculum.

2.1.2. Measures

A description of which measures were administered and when appears in Table 1.

2.1.2.1. School decile and curriculum. In NZ, the Ministry of Education calculates a decile rating for each school based on the proportion of students attending from low socio-economic status communities. The decile scores (each representing 10 percentiles) are integers ranging from one to 10, with ten indicating greater affluence. School curriculum indicated whether the school followed the state or Steiner curriculum.

2.1.2.2. Classroom instruction. In both 2007 and 2008, teachers were asked to complete a self-report questionnaire of classroom reading instruction practice. The questionnaire contained six domains derived from research on classroom activities associated with learning to read (Foorman & Schatschneider, 2003). The domains were: oral language, learning about books, learning about sounds, learning about meaning, reading of text, and writing. Teachers were asked to describe the kinds of activities engaged in during class time, providing an accompanying time

Table 1
Study design including measures, cohorts, and testing points.

Cohort	Testing point	Calendar year tested	State school measures	Steiner school measures
Junior-Elementary School		Year 1/Class I		
	1	2007, beginning	ISF, LNF, PPVT-IV	ISF, LNF, PSF, NWF, PPVT-IV
	2	2007, middle	ISF, LNF, PSF, NWF, HLEQ	ISF, LNF, PSF, NWF, HLEQ
	3	2007, end	LNF, PSF, NWF	LNF, PSF, NWF
		Year 2/Class II		
	4	2008, beginning	LNF, PSF, NWF	LNF, PSF, NWF
	5	2008, middle	PSF, NWF, ORF, RSC	PSF, NWF, ORF, RSC
	6	2008, end	PSF, NWF, ORF, RSC, WI, WA	PSF, NWF, ORF, RSC, WI, WA
Middle-Elementary School		Year 3/Class III		
	7	2007, beginning	NWF, ORF, PPVT-IV	NWF, ORF, PPVT-IV
	8	2007, middle	ORF, HLEQ	ORF, HLEQ
	9	2007, end	ORF	ORF
		Year 4/Class IV		
	10	2008, beginning	ORF	ORF
	11	2008, middle	ORF, RSC	ORF, RSC
	12	2008, end	ORF, RSC, WI, WA,	ORF, RSC, WI, WA
Senior-Elementary School		Year 5/Class V		
	13	2007, beginning	ORF, PPVT-IV	ORF, PPVT-IV
	14	2007, middle	ORF, HLEQ	ORF, HLEQ
	15	2007, end	ORF, WI, WA	ORF
		Year 6/Class VI		
	16	2008, beginning	ORF	ORF
	17	2008, middle	ORF, RSC	ORF, RSC
	18	2008, end	ORF, WI, WA, PC, RSC	ORF, WI, WA, PC, RSC

Note: ISF, initial sound fluency; LNF, letter naming fluency; PSF, phonemic segmentation fluency; NWF, nonsense word fluency; ORF, oral reading fluency; PPVT-IV, Peabody Picture Vocabulary Test IV; HLEQ, home literacy environment questionnaire; RSC, reading self-concept; WI, word identification; WA, word attack; PC, passage comprehension.

estimate. Although self-report measures may be subject to bias (e.g., response, desirability), teacher-report provides a broad perspective of classroom instruction from a social validity perspective (Erchul & Sheridan, 2008; Gersten, Baker, Haager, & Graves, 2005). Preliminary evaluation of this measure indicated it captured the transition from code-based to comprehension instruction across the grades (Barnes, 2008).

2.1.2.3. Peabody Picture Vocabulary Test IV (PPVT-IV). The PPVT-IV was administered in the middle of 2007 to assess receptive vocabulary, which is strongly linked to both reading and verbal skill (Hodapp & Gerken, 1999; Sénéchal, Ouellette, & Rodney, 2006). This measure (a) has good supporting technical adequacy evidence (Dunn & Dunn, 2007); (b) is widely used in research (Sénéchal et al., 2006); and (c) performs well in NZ (Cleveland & Reese, 2008; Reese, Jack, & White, 2010; Reese & Read, 2000). Standardized scores are presented. In some instances children were absent for initial PPVT testing, thus this was administered at the next possible testing point.

2.1.2.4. Home literacy environment questionnaire (HLEQ). The HLEQ was employed to account for differences in children's HLE and was adapted from a previously validated measure (Griffin & Morrison, 1997) to encompass the NZ state and Steiner contexts and increases in technology used in homes since the questionnaire's inception. There were nine literacy-related questions. Factor analyses on this measure indicate a four-factor solution comprising: parent-literacy activities, child-literacy activities, parent–child language activities, and general literacy environment (Schaughency, Suggate, & Reese, 2008). Data with middle elementary children suggest that the scores correlate with literacy attainment, and the subscales with one another (Schaughency et al., 2008), and a similar scale demonstrated similar properties in Germany (Niklas & Schneider, 2010). An overall composite score was used to reduce degrees of freedom in analyses. Interscorer agreement was estimated as 95 percent, calculated from 22 percent of the questionnaires. The return rates were 85 versus 76 percent for the later versus earlier RIA, respectively.

2.1.2.5. Demographic data. Ethnicity was collected from school archives, and five demographic questions added to the HLEQ provided data on parental education and occupation. School archives were checked for these data to replace missing data from non-returned HLEQs. Because of only partially complete archival data on whether children spoke a language other than English at home (hereafter: second language spoken at home), during the middle of the second year of the study, children were asked whether they spoke a second language at home. A correlation between children's report and school archives indicated excellent criterion validity ($r = .80$, $n = 117$, $p < .01$) so these scores were used as a proxy for children's exposure to languages other than English. Parental occupation was scored using the Elley–Irving Socio-Economic Index: 2001 Census Revision, with 6 indicating the occupations with the least status and 1 the highest, as reflected in the average income and education levels of workers (Elley & Irving, 2003). Parental education was scored as follows: 1 = No high school, 2 = three years of high school, 3 = four to five years of high school, 4 = obtained University Entrance requisite, 5 = post-school diploma or trade certificate, 6 = university graduate.

2.1.2.6. Reading self-concept. A three-item orally administered measure of children's reading self-concept was given at the end of 2008, because research suggests that older children are better able to estimate reading self-concept (Chapman & Tunmer, 1995). Questions related to three important constructs, namely perceived competence, difficulty, and attitudes toward reading (Chapman and Tunmer). Children's responses to these questions intercorrelated moderately to strongly and, particularly for older children, related to oral reading fluency (Schaughency, Suggate, Tustin, & Madigan, 2009). Responses were summed into a single index.

2.1.2.7. Reading measures. Woodcock Reading Mastery Test Revised – Normative Update. Three subtests of the Woodcock Reading Mastery Test Normative Update (WRMT-NU) (Woodcock, 1998) were administered (at the end of year/class 2, 4, 5, and 6), which have been used extensively in previous research (e.g., Hosp & Fuchs, 2005). These tests assessed word reading (Word Identification),

Please cite this article in press as: Suggate, S. P., et al. Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly* (2012), http://dx.doi.org/10.1016/j.ecresq.2012.04.004

non-word reading (i.e., Word Attack), and reading comprehension (i.e., Passage Comprehension). Alternate forms (Form G and H) were used so that children never received the same form twice. Raw scores are reported. Due to practical constraints, Passage Comprehension could only be administered to the older children at the end of the study. Importantly, factor analyses indicate that Passage Comprehension includes a strong decoding and comprehension component (Keenan, Betjemann, & Olson, 2008).

Concurrent validity coefficients for these measures are excellent for the US standardization sample, particularly for children around the age of those in this study (Woodcock, 1998). Direct empirical evaluations of the Woodcock tests in New Zealand are scarce, although the *Woodcock Johnson III* has been recently adopted in NZ by SPELD, a not-for-profit agency providing services to individuals with suspected learning disabilities and professional development to the education sector (Brooking & Hodgen, 2010). To empirically evaluate the appropriateness of WRMT-NU with New Zealand school children, we examined concurrent validity of the WRMT-NU subtests to be used in this study with independent samples of children who were in Years 2 ($n = 62$) and 3 ($n = 39$) in local state schools. In our Year 2 sample, we correlated the WRMT-NU subtests with the New Zealand Word Identification Fluency task (McLennan, Schaughency, & Struthers, 2011; McLennan, Schaughency, Struthers, & Abraham, 2011; McLennan, Struthers, Schaughency, & Clarke, 2010), developed through a curricular sampling approach (Fuchs, Fuchs, & Zumeta, 2008), sampling from words to which NZ children (in the state school curriculum) are exposed in their reading instruction, informed by research on the words that beginning NZ school children use in their writing (Randell, 2005). Correlations of the WRMT-NU subtests used in the study with NZWIF ranged from $r = .71$ for Word Attack, $r = .87$ for Passage Comprehension, to $r = .90$ for Word Identification, thereby suggesting that, for children in Year 2 at least, WRMT-NU performance is robustly correlated with proficiency in reading words encountered and used by NZ school children. In addition, at the end of the school year, children's teachers were asked to indicate whether or not they were concerned about participants' reading progress. We compared the concern and no concern groups in Years 2 and 3 on WRMT-NU subtest performance and found that in each comparison children in the no concern group outperformed those in concern group ($p < .01$), providing known groups validity evidence from a social validity perspective (Schaughency & Suggate, 2008). Thus, although developed in the US, available data provide preliminary support for the technical adequacy of the WRMT-NU in reading research in the New Zealand context.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS). Multiple antecedents of reading acquisition have been identified (Snow et al., 1998). Longitudinal research examining early predictors of later reading suggests the importance of incorporating measures across these domains (e.g., assessing both phonemic awareness and alphabetic knowledge) (Manolitsis, Georgiou, Stephenson, & Parilla, 2009) and using developmentally appropriate measures to tap into these domains (e.g., phonemic awareness) (Manolitsis, Georgiou, & Parilla, 2011). DIBELS are a series of brief and developmentally sequenced measures of early literacy skills (Kaminski, Cummings, Powell-Smith, & Good, 2008), focusing on the alphabetic and phonological principles for younger children and for older children's reading fluency. In NZ samples, DIBELS indices correlate well with reading across primary school (Schaughency & Suggate, 2008; Schaughency, Suggate, & Tustin, 2010; Struthers, Schaughency, Suggate, Clarke, & Thurlow, 2010). Here, the developmentally sequenced measures of phonemic awareness were Initial Sound Fluency, an early onset phoneme recognition task (de Graaf, Hasselman, Bosman, & Verhoeven, 2008; Manolitsis et al., 2011), and Phonemic Segmentation Fluency, involving developmentally more difficult skills of identifying onset, medial, and the final phonemes (de Graaf et al., 2008; de Graaf, Hasselman, Verhoeven, & Bosman, 2011). Letter Naming Fluency tapped early alphabet knowledge with Nonsense Word Fluency (NWF) and Oral Reading Fluency (ORF) assessing aspects of accuracy and fluency in decoding and reading of connected text. In a meta-analysis ORF exhibited strong correlations across grades 1–6 with measures of reading and reading comprehension (Reschly, Busch, Betts, Deno, & Long, 2009), supporting its use here. Scores on the measures represent the number of letters correctly named, phonemes correctly segmented, non-word letter sounds (NWF) or words (ORF) correctly read per minute. NWF consists of alternate forms of two (VC) or three (CVC) letter pseudowords during the period that it is administered. The difficulty of the ORF passages increases with grade, with simpler, shorter and more controlled texts in the beginning.

To adapt DIBELS to the Steiner samples and the study requirements, three alterations were made. First, US kindergarten-aged children at the beginning of school do not receive Phonemic Segmentation Fluency and NWF according to the DIBELS screening protocol; however, these were administered to investigate whether the junior-elementary school Steiner children, who would be similarly aged to children in their second or third year of schooling in the state school sample, had acquired decoding and phonemic segmentation skills prior to school entry. Second, because of the anticipated difficulty in completing NWF for children in at the beginning of school, a discontinue criterion was introduced (<5 on Letter Naming Fluency and no NWF practice items correct). Third, Steiner pupils were given one age-matched and one schooling-matched ORF probe. Thus, a Class IV student at a Steiner school, for example, was administered both a Year 4 probe consistent with his exposure to schooling, and a Year 5 probe, consistent with similarly aged children attending state schools.

We selected NWF and ORF as our primary longitudinal measures of decoding and reading fluency, based on the following rationale. Whereas the appropriateness of ORF as a measure of reading outcome has not been firmly established prior to about the middle of the second year of school (Wayman, Wallace, Wiley, Tichá, & Espin, 2007), NWF is closely related to reading of connected text in this age group (Baker, Park, & Baker, 2012). In NZ, NWF has a strong relationship with measures of word reading at the beginning of Year 2 (e.g., $r = .73$, McLennan et al., 2010) and actual reading of connected text at the end of Year 2 (e.g., $r = .70$, Suggate & Schaughency, 2007). Moreover, norms suggest that the transition in trajectories from NWF in year 1 and year 2 to ORF in year 3 is 'smooth'. In the current study, ORF and NWF scores at the end of year/class 2 correlated ($r = .71$, $p < .001$) and a paired samples t-test comparing NWF and ORF at this point found no significant difference, $t(120) = .71$, $p = .48$. Conversely, although NWF and ORF scores still correlate highly at the beginning of year/class 3 ($r = .74$, $p < .001$), the mean raw scores are no longer similar, $t(103) = 4.83$, $p < .001$, supporting the transition to ORF. Therefore, we modeled reading development in the first two years of school with NWF and thereafter with ORF.

2.1.3. Procedure

There were six testing points, three each in 2007 and three in 2008 for each cohort. The testing points occurred at the beginning, middle, and end of the school year. Across the junior, middle, and senior cohorts, six years of elementary school were studied, three times per year, giving six testing points per child and 18 across the entire elementary period. The measurement schedule and information of cohorts, testing points, ages, and placement in school (e.g., year 4) are depicted in Table 1. Due to child absences, seven late RIA and two early RIA children had the PPVT-IV administered up to 10 months later than indicated in Table 1 (5.71 and 5.00 months later on average, respectively).

Trained native English-speaking psychology-graduate students and one post-doctoral researcher administered the tasks.

Table 2
Teacher reported classroom reading-related activities as a function of school type.

Classroom activity	State school		Steiner school		*d*
	M	*SD*	*M*	*SD*	
Oral language					
Minutes	116.87	84.49	196.54	104.88	−.88[*]
Percentage	15.56	9.63	34.15	17.24	−1.54[*]
Learning about books					
Minutes	93.99	106.22	48.75	47.31	.48[a,*]
Percentage	10.65	8.55	7.61	5.22	.39
Learning about sounds					
Minutes	75.83	48.61	74.79	124.36	.01[a]
Percentage	9.99	6.07	9.35	8.98	.09
Learning about meaning					
Minutes	130.17	110.96	48.21	40.73	.84[a,*]
Percentage	15.28	8.66	7.78	6.28	.92[*]
Reading text					
Minutes	190.87	118.05	95.00	73.56	.88[*]
Percentage	24.05	10.7	15.15	9.48	.86[*]
Written language and spelling					
Minutes	176.33	63.34	151.49	72.87	.38
Percentage	24.47	7.88	25.96	12.57	−.16[a]
Text-related activities total					
Minutes	761.17	375.99	466.99	311.58	.82[*]
Percentage	84.44	9.63	65.85	17.24	1.54[*]

Note: Data are based on 39 state and 14 Steiner returned and valid teacher questionnaires.
[a] Levene's test for variance equality significant ($p < .05$), thus equal variance not assumed.
[*] $p < .05$.

Administrators were trained in accordance with the administration and scoring criteria and achieved interscorer reliability coefficients often in excess of .95 (minimum of .80) on five consecutive administrations with a speaker of NZ English.

Blind coding of half of the classroom instruction measures was conducted by the first author and an advanced undergraduate psychology student. This involved firstly a domain validity check to ensure that teachers had placed the particular activity in the correct category (e.g., was learning about the placement of a book's title on the title page categorized as learning about books activity or as reading text?). Because this resulted in researcher decisions about where to place activities, we calculated interscorer agreement on the amount of time recorded as spent by teachers in each of the categories. When below .80 for any given domain, the form was recoded, a procedure that resulted in initial disagreement for 43 percent of the forms, although the mean reliability across categories ($r = .86$ to .92), and forms ($r = .75$ to 1.00) was excellent. Finally, after independent recoding and scoring, the reliability criterion of .80 for each category was achieved on all but three occasions, with the estimated reliability coefficient being .91.

2.2. Results

2.2.1. Background variables

Analyses were firstly conducted to compare the earlier and later RIA samples on key background variables to explore the extent to which background factors played a role in between-group reading performance. These background factors were measured at the school, classroom, and child levels and are thus reported accordingly.

2.2.1.1. School-level variables. The mean weighted decile rating (which reflects SES), as a function of RIA, was compared with a between-subjects *t*-test with non-equal variances assumed (because Levene's test suggested unequal variances, $p < .05$). This indicated a slightly greater mean decile rating for the later RIA ($M = 6.61$, $SD = 1.01$) than the earlier ($M = 6.18$, $SD = 2.44$), $t(286) = 2.08$, $p < .05$, $d = .20$.

2.2.1.2. Classroom activities. The total amount of time reported to be spent engaging in the literacy activities on the classroom instruction measure was, per week, 13 hours and 4.05 minutes ($SD = 5$ hours and 11.60 minutes) in the state schools and 10 hours and 14.77 minutes ($SD = 5$ hours and 43.02 minutes) in the Steiner schools, a difference that was not significant, $t(51) = 1.70$, $p = .10$. The amount of time spent in the different domains of the reading instruction activities is reported in two different forms. In the first, we simply report the average number of minutes recorded in each activity, and in the second, we report the percentage of each teacher's times reported to be devoted to the activities as a function of the total time spent on reading for that teacher. We calculated these two measures because the raw minutes recorded as spent in reading time varied across teachers, presumably because some teachers did not take overlap in different domains into account when estimating times. Accordingly, the raw number of minutes tended to provide data with skewed distributions, so we preferred the percentage devoted to each activity as the key dependent variable of classroom instruction for the later multilevel linear model (MLM) analyses but report both in Table 2.

As can be seen in Table 2, the state school teachers reported spending a greater proportion and quantity of their time teaching children about books, learning about meaning, and reading connected text, with the Steiner teachers focusing more on oral language activities but having a similar focus on learning about the sounds in words and engaging in writing activities. The state school children spend more time overall in text-related reading instruction.

To explore how teachers' devotion of their time to the various literary activities varied across grades and school type, 2×3 MANOVAs were conducted. The dependent variable was percentage because these data satisfied the statistical assumption of being normally distributed. There were significant main effects for cohort, $F(2, 52) = 3.73$, $p = .03$, school curriculum, $F(1, 52) = 24.47$, $p < .001$, and their interaction for oral language activities, $F(2, 52) = 4.13$, $p = .02$. State-curriculum teachers reported spending more time teaching about meaning, $F(1, 52) = 10.08$, $p = .003$, and older cohorts engaged more in written activities, $F(2, 52) = 5.34$, $p = .008$.

Please cite this article in press as: Suggate, S. P., et al. Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly* (2012), http://dx.doi.org/10.1016/j.ecresq.2012.04.004

Table 3
Demographic characteristics of participants, by elementary level cohort and reading instructional age group.

Proportion	Junior cohort			Middle cohort			Senior cohort		
	Early RIA	Later RIA		Early RIA	Later RIA		Early RIA	Later RIA	
	% (*n*)	% (*n*)	χ^2	% (*n*)	% (*n*)	χ^2	% (*n*)	% (*n*)	χ^2
Sex									
Girls	40(23)	70(21)	7.47*	47(35)	47(15)	.00	50(34)	50(11)	.00
Second language spoken at home									
Yes	6(4)	23(7)	5.61*	3(2)	13(4)	4.10	8(5)	27(6)	6.00*
Ethnicity									
NZE	89(54)	70(21)	6.11[a,]*	89(67)	84(27)	2.32[a]	78(52)	96(21)	3.84[a]
NZM	6(4)	10(3)		8(6)	6(2)		10(7)	0(0)	
Other	5(3)	20(6)		3(2)	9(3)		12(8)	5(1)	

Note: RIA, reading instruction age; NZE, New Zealand European; NZM, New Zealand Māori.
[a] Cell count low for χ^2; thus interpret cautiously; due to rounding cell means may not add to 100. Second languages spoken at home were most often Māori, and less frequently Asian languages, German, or French. The "other" ethnicity category included mostly Chinese, non-Chinese Asians, Pacific Islanders, and continental Europeans.
* $p < .05$.

2.2.1.3. Child factors. Tables 3 and 4 present ethnicity, whether a second language was spoken at home, sex, PPVT-IV, parental education and occupation, reading self-concept, and HLEQ data. In Table 3 for the junior cohorts, there were fewer girls, second language speakers, and a less ethnically diverse sample for the early RIA children. The trends for the later RIA children in Table 4 indicated statistically significant advantages for those with a later RIA for: HLEQ and chronological age, but a disadvantage in the junior cohort for reading self-concept. Although there were no overall differences in PPVT-IV scores, the later-RIA junior cohort had statistically significantly greater scores. A 2 × 3 MANOVA was conducted to explore HLEQ subscale scores, as a function of cohort and RIA. No interactions were significant but early RIA children engaged in more independent reading activities throughout, $F(1, 220) = 4.05$, $p < .05$, watched more television, $F(1, 220) = 140.87$, $p = .00$, engaged in fewer parent–child literacy activities, $F(1, 220) = 7.76$, $p < .01$, and their parents spent less time reading, $F(1, 220) = 9.89$, $p < .01$.

Analyses conducted on the entire sample as a function of RIA suggested that later RIA cohorts: (a) contained more children speaking a second language at home, $\chi^2(1, 220) = 15.06$, $p < .05$, (b) were older, $t(287) = 4.73$, $p < .05$, (c) had a lower RSC, $t(287) = 2.50$, $p < .05$, (d) had mothers who earned less, $t(227) = 2.13$, $p < .05$, but fathers who earned more, $t(191) = 3.12$, $p < .05$, and who were better educated, $t(199) = 2.66$, $p < .05$, and (e) a better HLE, $t(226) = 5.94$, $p < .05$.

2.2.2. Reading skill outcome variables

Having explored between-group differences in school, classroom, and child factors we now turn to reading performance, measured at each of the time points.

2.2.2.1. Reading skill and time in school. To investigate how the decoding and reading fluency of the children developed as a function of RIA, descriptive and comparative statistics are presented in Table 5. As Table 5 shows, the earlier RIA sample had superior initial decoding and reading fluency skills, with this advantage decreasing across time in the junior cohort (testing points 1–6). Across middle elementary school, the initially non-significant advantage for

Table 4
Demographic characteristics of participants by elementary level cohort and reading instructional age group.

RIA	Junior cohort				Middle cohort				Senior cohort			
	Earlier	Later	*t*	*d*	Earlier	Later	*t*	*d*	Earlier	Later	*t*	*d*
Age	87.89	101.40	11.75*	−2.61	110.40	125.59	15.71*	−3.32	134.00	148.23	10.52*	−2.58
SD	(5.73)	(3.76)			(4.85)	(3.86)			(6.09)	(3.01)		
n	63	30			75	32			67	22		
Mat Ed.	3.58	3.74	.50	−.11	3.59	4.00	1.30	−.29	3.46	4.00	1.20	−.35
SD	(1.47)	(1.23)			(1.49)	(1.16)			(1.66)	(1.21)		
n	52	27			58	28			46	16		
Mat Occ.	3.91	4.32	.83	−.21	3.40	4.13	1.61	−.39	3.28	3.83	.93	−.29
SD	(1.96)	(1.99)			(1.88)	1.92			(1.95)	(1.40)		
n	54	22			63	24			54	12		
Pat Ed.	3.46	4.04	1.59	−.40	3.53	4.04	1.40	−.35	3.56	4.27	1.66	−.51
SD	(1.49)	(1.36)			(1.50)	(1.40)			(1.58)	(.70)		
n	46	23			53	23			41	15		
Pat Occ.	2.77	1.78	2.60*	.73	2.73	2.11	1.60	.44	2.98	2.43	.80	.32
SD	(1.54)	(.73)			(1.50)	(1.13)			(1.73)	(1.40)		
n	43	18			55	18			50	7		
PPVT-IV	109.71	100.13	3.50*	.78	106.60	106.91	.12	−.03	101.19	103.41	.68	−.17
SD	(11.66)	(13.65)			(11.33)	(12.54)			(14.46)	(8.75)		
n	63	30			75	32			67	22		
RSC	12.60	10.53	4.59*	1.02	11.72	11.36	.83	.17	11.81	12.50	1.60	−.39
SD	(2.07)	(1.93)			(2.14)	(1.88)			(1.83)	(1.54)		
n	63	30			75	32			67	22		
HLEQ	17.17	19.09	2.27*	−.54	16.73	20.21	4.82*	−1.10	14.74	18.28	3.14*	−.91
SD	(3.65)	(3.37)			(3.48)	(2.30)			(3.44)	(5.01)		
n	63	27			59	28			46	16		

Note: RIA, reading instruction age; Ed., education; Occ., occupation; Mat, maternal; Pat, paternal; PPVT-IV, Peabody Picture Vocabulary Test; RSC, reading self-concept; HLEQ, home literacy environment questionnaire.
* $p < .05$.

Table 5
Descriptive and comparative statistics for decoding (NWF) and reading fluency (ORF) scores at each testing point as a function of earlier and later readers.

Cohort	Testing point	Early RIA			Later RIA			*t*	*d*
		M	*SD*	*n*	*M*	*SD*	*n*		
Junior	1	49.40	53.22	5[a]	4.48	7.93	27	4.44*	2.16
	2	33.18	32.76	61	11.44	12.03	25	3.22*	.76
	3	40.62	29.33	61	12.11	10.63	28	4.98*	1.14
	4	46.13	30.38	61	14.30	10.74	27	5.29*	1.22
	5	57.75	31.68	59	24.96	12.86	26	5.08*	1.20
	6	73.19	47.88	59	38.03	15.66	29	3.84*	.87
Middle	7	86.32	45.10	73	71.13	49.44	31	1.53	.33
	8	89.69	40.60	74	81.72	47.97	29	.85	.19
	9	87.53	38.53	73	85.93	45.96	29	.18	.04
	10	100.69	38.70	72	99.42	48.39	31	.14	.03
	11	97.26	39.72	74	105.55	42.36	31	.96	−.20
	12	94.97	38.69	74	106.65	43.01	31	1.36	−.29
Senior	13	85.91	32.57	66	104.94	23.55	16	2.20*	−.61
	14	92.92	32.82	66	114.67	35.62	21	2.59*	−.65
	15	109.06	33.83	65	129.81	33.38	21	2.45*	−.61
	16	110.07	37.38	67	149.47	25.96	17	4.09*	−1.11
	17	105.45	35.64	67	128.17	29.60	18	2.48*	−.66
	18	107.83	34.83	63	124.42	25.50	21	2.02*	−.51

Note: NWF, nonsense word fluency (phases 1–6); ORF, oral reading fluency (phases 7–18).
[a] This sample comprises the 5 Waldorf pupils who had previously attended the state schools, thus having an RIA of 5 years.

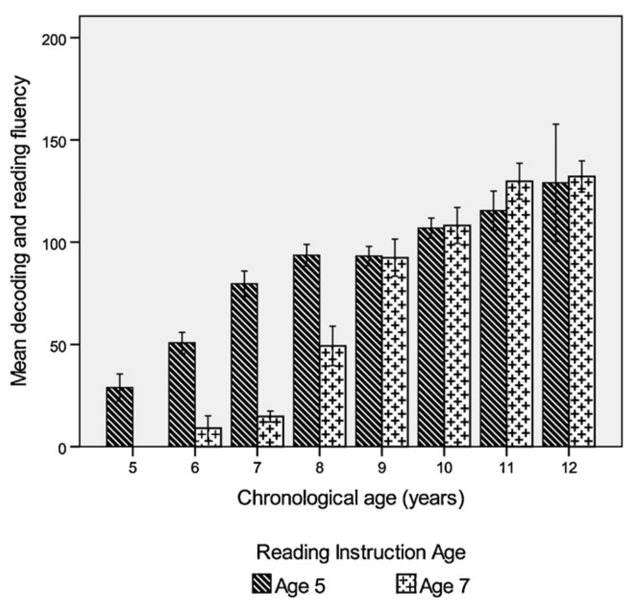

Fig. 1. Mean decoding and reading fluency scores as a function of chronological age for children with a reading instruction age of five or seven. Error bars represent two SEM.

the earlier RIA is reversed, with the advantage in these schooling-matched comparisons favoring the later RIA for the entire senior elementary-school period. To facilitate interpretation, these data are presented graphically in Fig. 1, however, this time as a function of age, not schooling.

2.2.2.2. Multilevel linear modeling (MLM). We next conducted a series of MLMs to investigate whether the pattern of findings in Fig. 1 changed once we controlled for school-, classroom-, and child-level factors. The covariate and reading-measures data did not exhibit significant skew and kurtosis. To facilitate interpretation of the model, all covariates (i.e., PPVT-IV, reading self-concept, decile, classroom instructional variables), except for chronological age and testing point, were grand-mean centered. An interaction between RIA and testing point (both linear and loglinear to test for non-linear growth) tested the idea of whether the trajectories as a function of RIA would converge or diverge. One further interaction between RIA and chronological age accounted for growth difference by age, independent of testing point.

In model building, the first priority was to account for as many potential confounding factors as possible, without compromising power. Therefore, MLMs were first constructed with the variables containing complete data, which necessitated initially excluding the HLEQ. Then, the HLEQ was added to the best model to see if this explained additional variance. On first inspection, the data would appear to be naturally nested into school, classroom, student, and time levels; however, because many children changed classrooms from one year to the next—and to account for differences in cohort in the sequential design—a cohort level instead of a classroom level was selected.

Unconditional model. The first step to assess the appropriateness of using MLM is to calculate a null model (Tabachnick & Fidell, 2007), here modeling decoding or reading fluency with random intercepts specified at three levels (i.e., time, student, cohort, and school). From this null model, the intraclass correlation coefficient suggested that growth over time accounted for 46.39 percent of variance in reading scores, student-level variables 19.82 percent, cohort 31.40 percent, and school-level variables 2.39 percent. The small amount of variance explained by the school level calls into question the need for a four-level model (and is probably due to the small number of schools in the study). Although this four-level model was better than a three-level model without school level, $\chi^2(1, N = 1750) = 3.88$, $p < .05$, when predictors were entered into the larger model, the Hessian matrix did not provide a solution with a positive random effect error term. Therefore, the three-level unconditional model was preferred and the school level was omitted and instead modeled with predictors at the subject level, as recommended by Nezlek (2010). In the three-level null model, the repeated measures, student, and cohort levels explained 46.44, 22.79, and 30.76 percent of the variance in decoding or reading fluency, respectively.

Full model. A full model was next constructed containing all predictors identified as important in the comparative analyses (Tables 2–5) and this improved the null model significantly, $\chi^2(13, N = 1722) = 379.85$, $p < .001$. Next, redundant predictors were trimmed from the full model (Tabachnick & Fidell, 2007), initially according to the more conservative criterion of $p < .20$ to

Table 6
Repeated-measures multilevel model with three levels predicting reading fluency.

	Estimate	Standard error	*t*
Intercept	−22.44	25.10	−.89
School-level covariate			
Decile	1.75	.83	2.12*
Classroom-level covariates			
Oral language instruction	.11	.07	1.60
Reading in class	.25	.10	2.64*
Student-level covariate			
Peabody Picture Vocabulary Test IV	.75	.14	5.25*
Reading self-concept	5.93	.87	6.79*
Student-level factors			
Early RIA	−9.47	31.22	−.30
Later RIA	–	–	–
Interactions			
RIA of 5 × Testing point	−1.56	1.33	−1.18
RIA of 7 × Testing point	4.08	1.49	2.75*
RIA of 5 × *ln*(testing point)	22.94	3.76	6.11*
RIA of 7 × *ln*(testing point)	.36	3.57	.10
RIA of 5 × Age(months)	.78	.30	2.59*
RIA of 7 × Age(months)	.56	.31	1.78†

Note: PPVT-IV, Peabody Picture Vocabulary Test IV; RIA, reading instruction age.
† $p<.10$.
* $p<.05$.

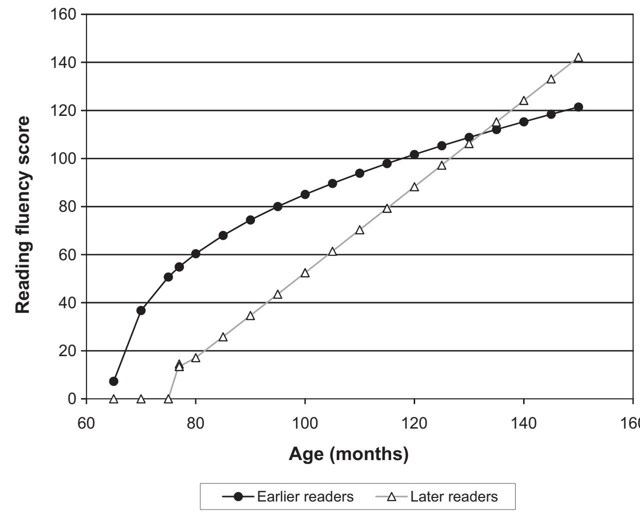

Fig. 2. Estimated decoding and reading fluency trajectories as a function of chronological age and reading instruction age from the multilevel linear model.

reduce the likelihood of erroneously deleting potentially meaningful covariates (Type II error). The variables of school curriculum ($p=.67$), sex ($p=.83$), second language ($p=.40$), and reported teaching about meaning ($p=.46$) were dropped. Even though the RIA dummy variable was not a significant predictor ($p=.60$), this variable was not excluded because of its centricity to the hypotheses. Removing these predictors did not result in an inferior model, $\chi^2(4,\ N=1722)=1.55$, $p=.82$. To investigate the effect of modeling reading growth logarithmically, the natural log of testing point was substituted, then combined with testing point in the reduced model. In both instances, the models containing both interaction terms were better than that without testing point, $\chi^2(2,\ N=1722)=8.82$, $p<.05$, and to that without the log of testing point, $\chi^2(2, N=1722)=35.38$, $p<.001$. Therefore, the resultant full model contained: school SES; oral language and reading time in class; PPVT-IV; reading self-concept; and the RIA by time, age, and the natural log of testing point.

The role of HLE and parental education and income. Having determined the optimal full model, the role of the home literacy environment was explored, but adding HLEQ did not improve the model fit, $\chi^2(1, N=1384)=.16$, $p=.69$, neither did adding parental income and education, $\chi^2(4, N=924)=6.71$, $p=.15$. Therefore, it is unlikely that RIA-related differences in HLE, as measured by the HLEQ, or parent education and income explained the findings.

The final model. The best model is presented in Table 6. In Table 6, the 'estimate' column indicates the extent to which an increase in one unit of the covariates (or level of the factors) increases the overall literacy outcome score. For example, children who were older by one month performed .55 points higher on the decoding and reading fluency tasks, and an increase in PPVT-IV score by one would, on average, suggest a higher reading score by .76 points. Thus, children with higher PPVT-IV and reading self-concept scores performed higher on the reading measures, and children's reading performance increased as they spent more time in school. There was a positive association between classroom reading instruction and oral language activities and reading, and a tendency toward higher reading achievement for higher decile schools.

Because of the logarithmic terms, the trajectories as a function of RIA are difficult to estimate, so the relationship is depicted in Fig. 2. Interestingly, this follows a similar pattern to the raw data modeled in Fig. 1. To determine the point at which the trajectories converged, the equations for reading as a function of RIA were solved. Thus letting the equation for the earlier RIA equal that of the later RIA gives:

$$-22.44-9.47+.78(\text{Age})+22.94\times\ln(T_{\text{E}})-1.56(T_{\text{E}})$$
$$=-22.44+.56(\text{Age})+.36\times\ln(T_{\text{W}})+4.08(T_{\text{W}})$$

where T_{E} is the testing point for earlier RIA and T_{L} is the testing point for later RIA. This equation solves to:

$$\text{Age}=130.69.$$

Therefore, from the above model, the point at which the later RIA reading trajectory meets the earlier RIA is 130.69 months, or 10.89 years.

2.2.2.3. Cross-sectional analyses.

Decoding and word reading. MANCOVAs were conducted to investigate reading achievement on the Woodcock measures of Word Identification and Word Attack and age level ORF as a function of RIA at specific points in school, because these data were not available longitudinally. As with the MLM, PPVT-IV, reading self-concept, chronological age, and school decile were included as covariates. For both 11-year-old, $F(3,\ 84)=.93$, $p=.43$, and 12-year-old children, $F(3,\ 74)=.05$, $p=.99$, the reading measures did not reveal any significant advantage attributable to RIA, extending the finding that the later RIA sample had achieved reading equivalence to measures additional to oral reading fluency (i.e., on Word Attack and Word Identification).

Reading comprehension. Passage Comprehension scores were only collected from the senior cohort at the end of the study. Age-norm equivalents were obtained to facilitate interpretation of comparisons across samples (because they had been given alternate forms as per Table 1). From these age-norm scores, chronological age was subtracted to reflect the number of months that children were reading above or below similarly aged children. Comparing children in the early RIA group ($M=2.63$ months below norm average, $SD=25.87$) with children in the later RIA group ($M=18.38$ months above norm average, $SD=79.28$), there was no statistically significant difference between these means, $t(82)=1.19$, $p=.25$, equal variance not assumed. However, two of the later RIA scores were clearly outliers (reading age equivalent of 16.67 and 23.92 years), which lead to a large standard deviation and positive skew for the later RIA sample. Therefore, comparing medians may instead provide the best comparative estimate of

Table 7
Scores for initial sound, letter naming, and phonemic segmentation fluency for children in their first year of school at the beginning of the study (i.e., junior cohort).

Testing point	Stat.	Initial sound fluency			Letter naming			Phonemic segmentation		
		RIA			RIA			RIA		
		Early	Later	*t*	Early	Later	*t*	Early	Later	*t*
Year/class 1										
1	*M*	17.59	20.89	−1.23	24.67	12.22	2.79*	–[a]	14.22	–
	SD	11.81	8.85		21.29	12.23		–[a]	15.15	
2	*M*	23.53	24.48	−.39	42.08	22.88	4.61*	18.85	16.80	.48
	SD	10.79	8.92		17.53	17.22		19.05	14.34	
3	*M*	–	–	–	50.36	29.07	5.93*	31.88	20.31	2.97*
	SD	–	–		15.70	16.09		18.21	14.93	
Year/class 2										
4	*M*	–	–	–	52.54	31.29	5.57*	35.36	24.50	2.81*
	SD	–	–		16.13	17.96		16.85	17.04	
5	*M*	–	–	–	–	–	–	40.98	34.54	1.68
	SD	–	–		–	–		16.70	15.37	
6	*M*	–	–	–	–	–	–	38.00	40.17	−.76
	SD	–	–		–	–		12.27	13.06	

Note: RIA, reading instruction age; Stat., statistic.
[a] Means not reported because only Steiner pupils were tested on Phonemic Segmentation Fluency at this phase.
* $p < .05$.

Passage Comprehension score. The medians indicated that, on average, the later RIA were performing about two months below US age equivalents, and the earlier RIA samples about nine months below.

Pre-reading measures. Before children can read text fluently, they experience development in a number of pre-reading domains (e.g., alphabet knowledge and phonemic awareness) (Snow et al., 1998). Therefore, we explored the development of these skills for the youngest children and these data are presented in Table 7. The younger and earlier RIA children exhibited superior letter knowledge for each comparison examining this variable, but this advantage was not found for initial comparisons on measures of phoneme awareness (i.e., Initial Sound Fluency and Phonemic Segmentation Fluency).

2.3. Discussion of Study 1

The data presented in Study 1 demonstrate that the later RIA children had no disadvantage at outcome in reading fluency, an important skill in reading achievement for this age of children (Hosp & Fuchs, 2005; Kaminski et al., 2008), by age 10.89 years, on average. These developmental trajectories suggest, therefore, that the later RIA generally caught up in reading to the early RIA children, even though the difference in RIA was estimated to be 19 months. Crucially, the beginning of school reading scores of the later RIA children suggested that many of these children had not yet developed more than rudimentary alphabetic knowledge after two months of schooling and therefore had likely not learned to read in the Steiner kindergarten setting.

The responses on the classroom instruction questionnaire suggested that the state school teachers did indeed report spending more time overall in school working with text whereas the Steiner children tended to be reported to receive more exposure to oral language activities that did not involve text. Accordingly, it does not appear that the Steiner children received intensive instruction in decoding-related skill per se to compensate for the later beginning; in fact, they may have had significantly less time devoted to developing decoding-related skills. Instead, they appear to have received more oral language experiences. The current findings point to the role of strong oral language activities in reading development in addition to decoding-related skills (e.g., NELP, 2008; Storch & Whitehurst, 2002).

Study 1 provides evidence of differential pre-reading skill development as a function of RIA. Specifically, the later RIA children entered school with very low levels of pseudoword decoding and letter naming skills, but their scores on phoneme awareness measures were comparable to the earlier RIA sample (when schooling-matched, not age-matched, comparisons were made). These data support hypotheses that phonemic awareness development to a point is not entirely dependent upon formal reading instruction and may develop implicitly through language (Walley, Metsala, & Garlock, 2003) and, given the age-matched superiority of the early RIA children's phonemic awareness, are also consistent with children experiencing a burst in phonemic awareness after reading instruction (Seymour et al., 2003).

The junior cohort also had statistically significantly different PPVT-IV scores, although both groups were still in the normal range. The causal reasons for this difference are unclear. Because this pattern was not observed on the other cohorts, it may be due to pre-existing differences between the groups; for example, the cohort with the advantage in PPVT-IV also had a greater proportion of girls. An alternative possibility is that the oral language-rich early school environment of the later RIA facilitated receptive vocabulary development of these students.

Interestingly, before removing outliers, the samples were performing above the US norms in reading comprehension at the end of elementary school and afterward the median was slightly below the US average. However, we urge that these US norm comparisons be interpreted cautiously given the difficulty in transferring norms from one country to another and comparing medians with normative means.

To consider reading development across the primary years, we followed three pairs of samples longitudinally, that is, across the first and second, third and fourth, and fifth and sixth years of elementary school, for children with an earlier versus later RIA. This design necessitates consideration of whether the data simulate those that would have been obtained from a purely longitudinal study. In general, evidence from the reading trajectories constructed across the cohorts suggested that these were smooth, such that the end development of one cohort seemed to plausibly map onto the beginning of the next, older cohort. An arguable exception to this occurred between the end of year/class 2 and the beginning of year/class 3 for the later-RIA children, with the reading fluency score increasing from 38.03 to 71.13 words per minute; however, this point represented the transition in measures from pseudoword reading to passage reading, appearing consistent with the nonlinear growth in Figs. 1 and 2 and those often observed during reading mastery (Paris, 2005). Moreover, NWF is expected to be more constrained, thus potentially suppressing these scores by the end of year/class 2. Therefore, a purely longitudinal design may have obtained similar results.

One limitation of Study 1, which led us to conduct Study 2, is the small sample size at the end of the study. Additionally, in this age group of children, ORF is a robust and sensitive indicator of reading performance, correlating highly with measures of reading comprehension (Hosp & Fuchs, 2005; Kaminski et al., 2008; Reschly et al., 2009). However, it could also be argued that reading fluency is a skill that may be less likely to reflect differences in vocabulary or comprehension that might be expected from the different early experiences of the groups. Therefore, we thought it important to extend the findings of Study 1 by looking more specifically at reading comprehension and vocabulary, well into formal schooling. Otherwise, we cannot discount the possibility that equifinality resulted because both groups of readers approached an asymptote in oral reading fluency, masking a difference in less constrained and important later reading skills, such as reading comprehension and vocabulary.

3. Study 2

In Study 2, we extend and replicate the findings of Study 1 in three important ways. First, we recruited additional data at the end point of the study on new samples of children because of concerns about the end-point sample size of the senior cohort. Second, we selected two samples that were more closely matched on age by recruiting Year 7 children in the state curriculum and comparing them with Class VI children in the Steiner curriculum. Accordingly, this provides a stronger test of the findings in Study 1 because the comparison involves older early RIA children. Third, we report more data on reading comprehension with these older children.

3.1. Method

3.1.1. Participants

All participants came from three schools with identical Ministry of Education decile ratings and located in small cities (population range 50,000–330,000). We excluded (a) one child for having a severe physical disability preventing reading task completion, (b) four children for whom complete reading data were not available, and (c) five children who began school early in the state curriculum, then transferred to the Steiner school, because this number comprised too few children to form a separate group. The early RIA participants ($n = 33$, $M = 12{:}6$ years, $SD = 3.50$ months) were on average six months older, $t(81) = 7.02$, $p = .00$, than the later RIA students ($n = 50$, $M = 12{:}0$ years, $SD = 4.47$ months) and participation rates were 55 and 33 percent, respectively. For early versus later RIA samples respectively, the samples were similar for (a) sex ratio, 46 versus 45 percent, respectively, $\chi^2(1, N = 83) = .00$, $p = .96$, (b) proportion of non NZ-European ethnic backgrounds, 11 versus 27 percent, $\chi^2(1, N = 82) = 1.91$, $p = .17$, and (c) for the proportion of children whose parents spoke another language at home, 9 versus 19 percent, $\chi^2(1, N = 68) = .86$, $p = .35$.

3.1.2. Measures

The measures used in Study 2 were similar to those used in Study 1. Background measures included again PPVT-IV, the HLEQ, parental occupation and education. The reading measures administered were Oral Reading Fluency, Word Attack, Word Identification, and Passage Comprehension (form G).

3.1.3. Procedure

The design of Study 2 was cross-sectional with testing occurring at the end of year 7 in the state schools and class VI in the Steiner school. The major differences in method to Study 1 were that (a) Passage Comprehension from Woodcock (1998) was administered to all children, (b) none of the DIBELS pre-reading measures were administered, (c) no classroom activity questionnaires were collected, and (d) reading self-concept data were not collected.

3.2. Results and discussion for Study 2

Twenty-nine of the early and 25 of later RIA families returned a completed HLEQ. Means and between-subject t-tests suggest that the early versus later RIA samples had, respectively, similar scores for (a) maternal profession, 3.33 versus 3.88, $t(67) = 1.28$, $p = .21$, (b) father's profession, 3.20 versus 2.81, $t(59) = .98$, $p = .33$, (c) father's education, 3.67 versus 4.10, $t(46) = 1.05$, $p = .30$, but not for (d) maternal education, 3.00 versus 3.79, $t(50) = 2.04$, $p < .05$, equivalent variance not assumed, or (e) HLEQ, 14.37 versus 16.85, $t(52) = 2.34$, $p = .02$. We investigated the subscale performance of the HLEQ using MANOVA and found that the only difference was that the earlier RIA children watched significantly more television, $F(1, 73) = 20.80$, $p = .00$.

Means, standard deviations, and independent samples t-tests were calculated to compare performance on the reading and vocabulary outcome measures as a function of RIA. The PPVT-IV scores of the samples were nearly identical, PPVT-IV, 107.92 versus 107.67, $t(81) = .09$, $p = .93$. On passage comprehension, the later RIA children performed significantly better $M = 46.39$, $SD = 7.70$, versus, $M = 42.82$, $SD = 7.83$, $t(81) = 2.05$, $p = .04$. Although in the same direction, none of the other means were significantly different. Thus: for word identification, $M = 79.10$, $SD = 10.21$, versus, $M = 81.36$, $SD = 10.21$, $t(81) = .98$, $p = .33$, for word attack, $M = 32.64$, $SD = 6.73$, versus, $M = 33.64$, $SD = 7.00$, $t(81) = .65$, $p = .52$, and for ORF, $M = 118.80$, $SD = 26.73$, versus, $M = 128.18$, $SD = 33.73$, $t(81) = .1.41$, $p = .16$.

The remaining analyses explored relations with passage comprehension, accounting for previously identified group differences. The focus was on passage comprehension as the outcome measure because of concerns that decoding and fluency represent constrained constructs (Paris, 2005), thus masking potential between-group differences, as groups approach ceiling on decoding and word recognition skills. In the first analysis, passage comprehension was predicted from RIA, chronological age, and PPVT-IV. PPVT-IV was included because of its strong ties to reading comprehension at this age (Cunningham & Stanovich, 1997). The overall model was significant, $R^2 = .56$, $F(3, 79) = 33.28$, $p = .00$, with RIA, $\beta = .29$, $p = .00$, and PPVT-IV, $\beta = .70$, $p = .00$, being significant predictors, but not chronological age, $\beta = -.10$, $p = .31$.

Finally, another regression analysis was conducted with RIA and PPVT-IV in the first step and HLEQ and maternal education in the second step. Chronological age was excluded because it was not a significant predictor in the earlier model. Maternal education did not explain unique variance but, due to missing maternal education data, did reduce the degrees of freedom in comparison to the model without it. Therefore, the regression analysis was conducted excluding maternal education. Adding HLEQ improved the model, $\Delta R^2 = .07$, $p = .02$, such that the final model was statistically significant, $R^2 = .61$, $F(3, 50) = 29.07$, $p = .00$, and PPVT-IV, $\beta = .63$, $p = .00$, and HLEQ were significant predictors, $\beta = .29$, $p = .00$, but RIA was not, $\beta = .18$, $p = .05$, missing the conventional cut-off value by $p = .001$. The likely reason for RIA being narrowly non-significant is due to the reduced sample size in the analyses with HLEQ.

The findings of Study 2 are consistent with Study 1 in that the reading achievement of children with a later RIA was similar to children with the early RIA. Importantly, Study 2 extended the findings of Study 1 by providing a better test of a conceptually less constrained construct; namely, reading comprehension (Paris, 2005). On this measure, the children with the later RIA achieved higher scores, and when controlling for HLE, this was significant at the level of $p = .05$.

Please cite this article in press as: Suggate, S. P., et al. Children learning to read later catch up to children reading earlier. *Early Childhood Research Quarterly* (2012), http://dx.doi.org/10.1016/j.ecresq.2012.04.004

4. General discussion

We presented data, conducted on English-speaking samples, demonstrating for the first time that children who begin learning decoding-related skills some 19 months later than their peers can eventually achieve equally in reading fluency – or even a little better on reading comprehension. Our findings suggest that success at reading is not assured by an earlier beginning. Although similar findings have previously been demonstrated internationally (Suggate, 2009), and for smaller RIA differences for kindergarten and preschool children (Durkin, 1974–1975; Schmerkotte, 1978), to our knowledge these data are the first obtained from English-speaking samples – a language renowned for the problems that it poses beginning readers (Seymour et al., 2003). Furthermore, our study mapped reading development across the entire elementary school period. These strengths of the current research temper findings from research investigating only short-term effects, and underscore the need for future research to look long-term.

With hindsight, we see the greatest weakness in our selection of child measures as being our failure to look at broader language measures other than receptive vocabulary and early phonemic awareness. Accordingly, we urge future research to consider expressive language, listening comprehension, and morphological awareness in addition to reading comprehension and vocabulary. Research might also look more in depth at motivational and patterns of reading behavior in later development, to more thoroughly explore possible benefits of early versus later reading. In addition, research might include a broader array of reading and educational outcome measures to discern whether there are differences in growth trajectories on learning within content areas. Although it would have been ideal to have explored such factors in this study as well, we considered it necessary that the first step in exploring the role of RIA start with more purely reading outcomes.

To our knowledge, a theoretical explanation for why this earlier advantage in reading would not be maintained has not been published and it would be premature to attempt one here. We note that our findings are consistent with the observations of the late Jeanne Chall (1976) who commented on the persistent finding that earlier advantages do not seem to hold up. This observation led her to look toward developing an understanding of developmental changes in the reading process over time (Chall, 1976). More generally, our findings may also be seen as consistent with the developmental principle of equifinality. This principle notes that there are often many different developmental pathways to the same outcome (Gottlieb, 2003), which, given the initial superiority of the early RIA children's decoding-related skills, was certainly observed here. Next we turn to consider alternative explanations for obtained results because the design of the study did not involve random assignment.

Three key factors exist that may account for why children with the later RIA attained equal, or slightly greater, reading achievement, namely those related to curricula, sample differences, and our measures. First, it is possible that methods of teaching reading in the Steiner curriculum are more effective than those used in the state curriculum, perhaps allowing the later RIA Steiner children to catch up in reading. The state curriculum in NZ has been criticized, and defended, for adopting an overly whole-word approach and neglecting explicit teaching of phonics (Thompson, 2003, but see Tunmer, Chapman, & Prochnow, 2004). However, this criticism could also be leveled at the Steiner curriculum, as in NZ both curricula have some similarity in the extent to which systematic phonics approaches were adopted. Moreover, the Steiner schools did not appear to compensate for the later RIA by offering more text-related instruction; on the contrary, they focused less on text-related skills across the first six years of school.

A second factor in curricula differences is the rich language approach in the later RIA schools. It is becoming recognized that the foundation of later reading, and in particular reading comprehension, is language (e.g., Verhoeven & Perfetti, 2011). Instead of focusing on developing decoding-related skills between the ages of five and seven, and in the first years of school, it may be that the environments in the Steiner kindergartens favored language development, which later feeds into reading comprehension (Cunningham & Stanovich, 1997; Sénéchal et al., 2006; Storch & Whitehurst, 2002). Indeed, the later-RIA children in the junior cohort did have a statistically significant advantage in receptive vocabulary scores, although this was not seen in the other cohorts. We suggest that future research explore receptive vocabulary development as a function of whether children are in early reading and/or language-rich environments, particularly as reading increases vocabulary in older children at least (Swanborn & de Glopper, 1999). Interestingly, we previously investigated the oral narrative skills, which tap semantic aspects of language, of children from Steiner kindergartens in comparison to state school children and did not find significant differences (Suggate et al., 2011).

The school experiences of participants in this study likely varied markedly between five and seven as a function of RIA. To measure classroom instruction, we used a self-report measure, whereas classroom observations may have provided a better measure of actual classroom practice. Thus, we also cannot rule out response bias without a direct observational measure. In defense of our measure, the responses from this questionnaire were consistent with descriptions of the respective educational philosophies and did contribute meaningfully to the MLM, such that school curriculum was a redundant predictor.

In terms of sample differences we found only limited evidence of differences in the non-reading covariates between samples and, when present, the advantages were not systematically in favor of either the earlier or later RIA, with the exception of the HLEQ and to a lesser extent whether a second language was spoken at home. In both studies, the earlier RIA children appeared to spend more time viewing television, and in the first study, although engaging more in independent reading, their parents were a little less involved in literacy activities. The relations between television viewing and reading achievement are complicated and the direction unclear (Beentjes & Van der Voort, 1988; Ennemoser & Schneider, 2007), although book reading certainly relates to greater reading achievement (Cunningham & Stanovich, 1997). Thus, the earlier RIA sample had a clear advantage in reading practice, with a potential disadvantage in parental involvement. There is therefore no clear indication from including the HLEQ in analyses that the later RIA group was able to catch up due to some particular feature of the home environment. Moreover, the parents of the later-RIA children had made an educational decision to opt out of the state system and enter their children in the Steiner schools, which likely suggests that, on average, their attitudes to education are different. Future research might consider how such attitudes might affect reading development.

Interestingly, there were also differences in the participation rates (in Study 1, 83.72% vs. 59.77%, in Study 2, 55% vs. 33%), with a greater proportion of the later RIA samples participating in the research in both studies. From our observations, we noted that participation rates seemed to be affected by a number of factors, including teacher interest in the study, size of the school, age of the sample, and our recruitment attempts. Additionally, because of the scarcity of the later RIA samples, we tried to ensure that all of these parents answered the questionnaire. A final consideration is the possibility that the interaction between the potentially increased parental involvement of the later RIA parents and the Steiner curriculum provides a rich base of language allowing children in these environments to offset early disadvantage in decoding-related skills and develop a slight edge in later reading comprehension. To explore the possibility that facilitation of language development via parental scaffolding (e.g., Reese, 1995) and/or the Steiner educational context might be the mechanism by which older RIA children develop the foundation for reading, future research might directly observe parent–child interactions and then compare later reading achievement as a function of RIA (earlier vs. later), curriculum (state vs. Steiner), and the quality of parental involvement (responsive scaffolding vs. less responsive or disengaged parent–child interactions; Davis, Evans, & Reynolds, 2010; Silinskas, Leppänen, Aunola, Parrila, & Nurmi, 2010).

5. Cautions and conclusions

We believe that understanding contextual influences on language development has practical implications for *how* educational systems, both in NZ and in other English-speaking countries such as the US and the UK, should foster literacy in the early years of schooling. However, we do not interpret the current findings as evidence that no reading instruction should occur before age seven for several reasons. First, our findings support the role of developing strong oral language skills, including phonemic awareness, as a foundation for reading, and these skills are today viewed as important aspects of pre-reading programs (e.g., Justice & Pullen, 2003). Second, we further point out that many children in the later RIA sample were six and a half, not seven years old when they began learning decoding-related skills. Third, our findings were not specific to disadvantaged children. We would describe both of our samples as, on average, having parents engaged in their children's education. Therefore, we cannot generalize these findings to other, less fortunate, contexts. In short, and as stated at the outset, we view this study as an important first step in investigating the complicated question of the effects of early reading instruction and acquisition of decoding-related skill on later skilled reading.

Until the above issues are resolved, we recommend that the findings of the present studies, which were necessarily non-experimental, be a stimulus for future research. One tantalizing possibility for this research is to verify whether the counter-intuitive finding—that equal, or even greater, long-term reading achievement can result despite delaying reading instruction by nearly two years—holds for (a) more phonics-based curricula, (b) at-risk samples, (c) families with a greater variation in HLE quality, and (d) whether effects are moderated by factors such as children's skill development. The implications that such research may have for conceptualizations of the formal and informal learning environments of preschool and kindergarten aged children are simply too great not to pursue.

Acknowledgments

This research was supported in part by research grants to the second author and a publishing bursary to the first author from the University of Otago, and an Alexander-von-Humboldt research fellowship to the first author. We would like to thank the following for assistance with this study: Sarah Fahey, Fiona Jack, Alana Roughan, Clark Sim, Tamara Suggate, Karen Tustin, and the participating school communities. Data analyzed to provide concurrent validity evidence for the *Woodcock Reading Mastery Test – Normative Update* were collected in the context of the ongoing Ph.D. research of Philippa Struthers in the Department of Psychology at the University of Otago. We gratefully acknowledge her work in sharing these data with us for these analyses.

References

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

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in Spanish and English. *Reading and Writing: An Interdisciplinary Journal*, *25*, 251–281. http://dx.doi.org/10.1007/s11145-11010-19261-z

Barnes, K. (2008). *Instructional content of New Zealand primary school curriculum as reflected by teacher-reported instructional practices*. Unpublished dissertation. Dunedin, New Zealand: University of Otago.

Beentjes, J. W. J., & Van der Voort, T. H. A. (1988). Television's impact on children's reading skills: A review of research. *Reading Research Quarterly*, *23*, 389–413.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson, & S. B. Newman (Eds.), *Handbook of early literacy research* (pp. 41–51). New York, NY: Guilford Press.

Boyce, L., & Snow, C. (2009). Introduction to the special issue. *Scientific Studies of Reading*, *13*, 95–98.

Brooking, K., & Hodgen, E. (2010). *Report of the Phase Two survey data prepared for SPELD NZ*. Wellington, New Zealand: New Zealand Council for Educational Research.

Burnett, J. (2007). A pleasurable path to literacy: Can Steiner contribute to the literacy debate? *Journal of Early Childhood Literacy*, *7*, 321–331. http://dx.doi.org/10.1177/1468798407083663

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

Cambridge Primary Review. (2009). *The final report. Cambridge Primary Review briefings*. Cambridge, UK: University of Cambridge.

Chall, J. S. (1987). Reading development in adults. *Annals of Dyslexia*, *37*, 240–251.

Chall, J. (1976, April). The great debate: Ten years later, with a modest proposal for reading stages. In *Paper presented at the conference on theory and practice of beginning reading instruction* Pittsburgh, PA,. Retrieved from:. http://eric.ed.gov/PDFS/ED155617.pdf

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationships with reading achievement. *Journal of Educational Psychology*, *87*, 154–167.

Chiu, M. M., & McBride-Chang, C. (2006). Gender, context, and reading. *Scientific Studies of Reading*, *10*, 331–362.

Clay, M. (1998). *An observation survey of early literacy achievement*. Auckland, New Zealand: Heinemann.

Cleveland, E. S., & Reese, E. (2008). Children remember early childhood: Long-term recall across the offset of childhood amnesia. *Applied Cognitive Psychology*, *22*, 127–142.

Cunningham, A. J., & Carroll, J. M. (2011a). The development of early literacy in Steiner- and standard-educated children. *British Journal of Educational Psychology*, *83*, 475–490. http://dx.doi.org/10.1348/000709910X522474

Cunningham, A. J., & Carroll, J. M. (2011b). Reading-related skills in earlier and later schooled children. *Scientific Studies of Reading*, *15*, 244–266.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934–945.

Davis, B. J., Evans, M. A., & Reynolds, K. P. (2010). Child miscues and parental feedback during shared alphabet book reading and relations with child literacy skills. *Scientific Studies of Reading*, *14*, 341–364.

de Graaf, S., Hasselman, F., Bosman, A. M. T., & Verhoeven, L. (2008). Cognitive and linguistic constraints on phoneme isolation in Dutch kindergarteners. *Learning and Instruction*, *18*, 391–403.

de Graaf, S., Hasselman, F., Verhoeven, L., & Bosman, A. M. T. (2011). Phonemic awareness in Dutch kindergarteners: Effects of task, phoneme position, and phoneme class. *Learning and Instruction*, *21*, 163–173.

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, *39*, 305–310.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*, 465–481.

Dockrell, J. E., Stuart, M., & King, D. (2010). Supporting early oral language skills for English language learners in inner city preschool provision. *British Journal of Educational Psychology*, *80*, 497–515. http://dx.doi.org/10.1348/000709910X493080

Dunn, L., & Dunn, D. (2007). *PPVT-IV, Peabody Picture Vocabulary Test Manual* (4th ed.). Minneapolis, MN: Pearson.

Durkin, D. (1974–1975). A six year study of children who learned to read in school at the age of four. *Reading Research Quarterly*, *10*, 9–61.

Edwards, C. P. (2007). Three approaches from Europe: Steiner, Montessori, and Reggio Emilia. *Early Childhood Research & Practice*, *4*(1).

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*, 250–287.

Elley, W. B. (1992). *How in the world do students read? IEA study of reading literacy*. The Hague, Netherlands: International Association for the Evaluation of Educational Achievement.

Elley, W. B., & Irving, J. S. (2003). The Elley–Irving socio-economic index: 2001 census revision. *New Zealand Journal of Educational Studies*, *38*, 3–17.

Ennemoser, M., & Schneider, W. (2007). Relations of television viewing and reading: Findings from a 4-year longitudinal study. *Journal of Educational Psychology*, *99*, 349–368.

Erchul, W. P., & Sheridan, S. M. (Eds.). (2008). *Handbook of research in school consultation*. New York, NY: Taylor and Francis.

Feitelson, D., Tehori, B. Z., & Levinberg-Green, D. (1982). How effective is early instruction in reading? Experimental evidence. *Merrill-Palmer Quarterly*, *28*, 485–494.

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn, & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning*. Baltimore, MD: Brookes.

Fuchs, L. S., Fuchs, D., & Zumeta, R. O. (2008). A curricular sampling approach to progress monitoring mathematics concepts and applications. *Assessment for Effective Intervention*, *33*, 225–233.

Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting reading outcomes for first-grade English learners: An observational study. *Remedial & Special Education*, *26*, 197–206.

Gottlieb, G. (2003). Probabilistic epigenesis of development. In J. Valsiner, & K. J. Connelly (Eds.), *Handbook of developmental psychology* (pp. 3–17). London, UK: Sage.

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

Greaney, K. (2002). Commentary: That reading debate again!. *New Zealand Journal of Educational Studies*, *37*, 101–108.

Griffin, E. A., & Morrison, F. J. (1997). The unique contribution of home literacy environment to differences in early literacy skills. *Early Childhood Development and Care*, *127–128*, 233–243.

Hodapp, A. F., & Gerken, K. C. (1999). Correlations between scores for Peabody Picture Vocabulary Test-III and the Wechsler Intelligence Scale for Children-III. *Psychological Reports*, *84*, 1139–1142.

Hosp, M., & Fuchs, L. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with age? *School Psychology Review*, *34*, 9–26.

House of Commons Education and Skills Committee. (2005). *Teaching children to read. Eighth report of session 2004–2005*. London, UK: The Stationery Office.

Justice, L. M., & Pullen, P. C. (2003). Promising interventions for promoting early literacy skills: Three evidence based approaches. *Topics in Early Childhood Special Education*, *23*, 99–113.

Kaminski, R. A., Cummings, K. D., Powell-Smith, K. A., & Good, R. H. (2008). Best practices in using Dynamic Indicators of Basic Early Literacy Skills for formative assessment and evaluation. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (pp. 1181–1204). Bethesda, MD: National Association of School Psychologists.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*, 281–300.

Leppänen, U., Aunola, K., Niemi, P., & Nurmi, J. E. (2008). Letter knowledge predicts Grade 4 reading fluency and comprehension. *Learning and Instruction*, *18*, 548–564.

Lillard, A., & Else-Quest, N. (2006). Evaluating Montessori education. *Science*, *313*, 1893–1894.

Lonigan, C. J., Schatschneider, C., Westberg, L., & The National Early Literacy Panel. (2008). *Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling (Developing early literacy: A scientific synthesis of early literacy development and implications for intervention)*. Jessup, MD: National Institute for Literacy & The Partnership for Reading.

Manolitsis, G., Georgiou, G. K., & Parilla, R. (2011). Revisiting the home literacy model of reading development in an orthographically consistent language. *Learning and Instruction*, *21*, 496–505.

Manolitsis, G., Georgiou, G. K., Stephenson, K., & Parilla, R. (2009). Beginning to read across languages varying in orthographic consistency: Comparing the effects of non-cognitive and cognitive predictors. *Learning and Instruction*, *19*, 466–480.

McLennan, K., Schaughency, E., & Struthers, P. (2011, June). Capturing transitions in human development: Evaluating an indicator of beginning reading in NZ across the second year of school. In *Paper presented at the 17th biennial AHDA conference* Dunedin, New Zealand.

McLennan, K., Schaughency, E., Struthers, P., & Abraham, R. (2011, November). Better serving all children in beginning reading instruction: Preliminary evaluation of a tool for indexing growth during reading acquisition. In *Presentation at the fourth educational psychology forum* Albany, New Zealand.

McLennan, K., Struthers, P., Schaughency, E., & Clarke, H. (2010, November). Towards evidence-based practice: Preliminary evaluation of a tool for monitoring progress in beginning reading in NZ. In *Presentation at the third educational psychology forum* Christchurch, New Zealand.

McNaughton, S., Phillips, G., & MacDonald, S. (2000). Curriculum channels and literacy development over the first year of instruction. *New Zealand Journal of Educational Studies*, *35*, 49–59.

Ministry of Education. (1999). *Literacy experts group report to the Secretary of Education*. Wellington, New Zealand: Ministry of Education.

Ministry of Education. (2003a). *Effective literacy practice in years 1–4*. Wellington, New Zealand: Learning Media.

Ministry of Education. (2003b). *Ready to read teacher support material. Sound sense: Phonics and phonological awareness*. Wellington, New Zealand: Learning Media.

Ministry of Education. (2007). *Ready to read teacher support material. Emergent level*. Author. Retrieved from http://www.tki.org.nz/r/literacy_numeracy/professional/teachers_notes/ready_to_read/literacy/index_e.php.

Ministry of Education. (2009). *The New Zealand curriculum reading and writing standards for years 1 to 8*. Wellington, New Zealand: Learning Media.

Ministry of Education. (2010). *The literacy learning progressions: Meeting the reading and writing demands of the curriculum*. Wellington, New Zealand: Learning Media.

National Early Literacy Panel (Ed.). (2008). *Developing early literacy: A scientific synthesis of early literacy development and implications for intervention*. Jessup, MD: National Institute for Literacy & The Partnership for Reading.

National Reading Panel (Ed.). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Bethesda, MD: National Institute of Child Health and Development. Retrieved from. http://www.nichd.nih.gov/publications/nrp/report.htm

Neuman, S. B. (2006). The knowledge gap: Implications for early education. In D. K. Dickinson, & S. B. Neuman (Eds.), *Handbook of early literacy research* (pp. 29–40). New York, NY: The Guilford Press.

Nezlek, J. B. (2010, March). *Workshop on multilevel analysis building a multilevel model*. Dunedin, New Zealand: Department of Psychology, University of Otago.

Ng, L. (2006). *Annual monitoring of reading recovery: The data for 2004*. Wellington, New Zealand: Ministry of Education.

Niklas, F., & Schneider, W. (2010). Der Zusammenhang von familiärer Lernumweltmitschulrelevanten Kompetenzen imVorschulalter [The interrelation of home literacy environment and different measures of performance at pre-school age]. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *30*, 149–165.

Olson, R. K., & Gayan, J. (2001). Brains, genes, and environment in reading development. In S. B. Neuman, & D. K. Dickinson (Eds.), *The handbook of early literacy research* (pp. 81–94). New York, NY: Guilford.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*, 184–202.

Quick, S., Lambley, C., Newcombe, E., & Aubrey, C. (2002). *Implementing the foundation stage in reception classes. Research report no. 350*. London, UK: The Stationery Office.

Randell, B. (2005). *PM benchmark kits: Reading levels and reading ages*. Melbourne, Australia: Thomson Learning.

Reese, E. (1995). Predicting children's literacy from mother–child conversations. *Cognitive Development*, *10*, 381–405.

Reese, E., Jack, F., & White, N. (2010). Origins of adolescents' autobiographical memories. *Cognitive Development*, *25*, 352–367.

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language*, *27*, 255–266.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement of oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*, 427–469.

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading, and spelling as predictors of school achievement and secondary school choice. *Learning and Instruction*, *18*, 201–210.

Schaughency, E., & Suggate, S. (2008). Measuring basic early literacy skills amongst year 1 students in New Zealand. *New Zealand Journal of Educational Studies*, *43*, 85–106.

Schaughency, E., Suggate, S., & Reese, E. (2008, July). Measuring the home literacy environment in a sample of New Zealand primary students. In *Paper presented at the society for the scientific study of reading* Asheville, NC,

Schaughency, E., Suggate, S., & Tustin, K. (2010, April). Using local norms to describe reading achievement: A longitudinal examination from third to fifth grade. In *Paper presented at the American education research association* Denver, CO.

Schaughency, E., Suggate, S., Tustin, K., & Madigan, S. (2009, December). Preliminary evaluation of brief self-ratings of reading: Do relationships with reading performance change with age? In *Paper presented at the educational psychology forum* Wellington, New Zealand.

Schmerkotte, H. (1978). Ergebnisse eines Vergleichs von Modellkindergarten und Vorklassen in Nordrhein-Westfalen [Results from a comparison of typical kindergartens and preschools in Northrhein-Westphalia]. *Bildung und Erziehung*, *31*, 401–411.

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson, & S. B. Neuman (Eds.), *Handbook of early literacy* (pp. 173–182). New York, NY: Guilford.

Seymour, P. H. K., Aro, M., Erskine, J. M., & COST Action Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Shanahan, T., & Lonigan, C. J. (2010). The National Early Literacy Panel: A summary of the process and the report. *Educational Researcher*, *39*, 279–285.

Silinskas, G., Leppänen, U., Aunola, K., Parrila, R., & Nurmi, J. E. (2010). Predictors of mothers' and fathers' teaching of reading and mathematics in kindergarten and Grade 1. *Learning and Instruction*, *20*, 61–71.

Smith, J. W., & Elley, W. (1997). *How children learn to read: Insights from the New Zealand experience*. Auckland, New Zealand: Addison Wesley Longman.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Steiner, R. (1924). *Conversations with teachers*. London, UK: Anthroposophic Press.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934–947.

Struthers, P., Schaughency, E., Suggate, S., Clarke, K., & Thurlow, J. (2010, April). Using a back-mapping framework to examine early literacy skills in the first year of school. In *Paper presented at the American education research association* Denver, CO.

Suggate, S. P. (2009). School entry age and reading achievement in the 2006 Programme for International Student Assessment (PISA). *International Journal of Educational Research*, *48*, 151–161.

Suggate, S. P. (2010). Why "what" we teach depends on "when": Grade and reading intervention modality moderate effect size. *Developmental Psychology*, *46*, 1556–1579.

Suggate, S. P., & Schaughency, E. A. (2007). The development of reading in Year 2 and its relation to reading in Year 3. *The Bulletin*, *109*, 40–42.

Suggate, S. P., Schaughency, E. A., & Reese, E. (2011). The contribution of age and formal schooling to oral narrative and pre-reading skills. *First Language*, *31*, 379–403. http://dx.doi.org/10.1177/0142723710395165

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Reviews of Educational Research*, *69*, 261–285.

Tabachnick, B. C., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Thompson, G. B. (2003). Teaching and the phonics debate: What can we learn? *New Zealand Annual Review of Education*, *11*, 161–178.

Tunmer, W. E., Chapman, J. W., & Prochnow, J. E. (2004). Why the reading achievement gap in New Zealand won't go away: Evidence from the PIRLS 2001 International Study of Reading Achievement. *New Zealand Journal of Educational Studies*, *39*, 127–145.

Verhoeven, L., & Perfetti, C. A. (2011). Introduction to this special issue: Vocabulary growth and reading. *Scientific Studies of Reading*, *15*, 1–7.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading & Writing: An Interdisciplinary Journal*, *16*, 5–20.

Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum based measurement in reading. *Journal of Special Education*, *41*, 85–120.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Newman, & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 11–29). New York, NY: Guilford.

Woodcock, R. W. (1998). *Woodcock Reading Mastery Tests – Revised Normative Update forms G and H examiner's manual*. Minneapolis, MN: Pearson.