



HAL
open science

Strict phonics beats mixed phonics: Effective teaching improves reading acquisition and reduces social inequalities

Johannes Ziegler, Paul Gioia, Jerome Deauvieu

► To cite this version:

Johannes Ziegler, Paul Gioia, Jerome Deauvieu. Strict phonics beats mixed phonics: Effective teaching improves reading acquisition and reduces social inequalities. 2024. hal-04769557

HAL Id: hal-04769557

<https://amu.hal.science/hal-04769557v1>

Preprint submitted on 6 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Strict phonics beats mixed phonics: Effective teaching improves reading acquisition and reduces social inequalities

Johannes Ziegler

Johannes.Ziegler@univ-amu.fr

Aix-Marseille Univ and CNRS <https://orcid.org/0000-0002-2061-5729>

Paul Gioia

École normale supérieure (ENS), Paris Science & Lettres (PSL), Centre Maurice Halbwachs (UMR 8097 CNRS/ENS/EHESS)

Jerome Deauvieu

École normale supérieure (ENS), Paris Science & Lettres (PSL), Centre Maurice Halbwachs (UMR 8097 CNRS/ENS/EHESS) <https://orcid.org/0009-0003-0483-9524>

Article

Keywords:

Posted Date: October 1st, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5017972/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Abstract

Reading is crucial for functioning in our modern societies, and reading failure is associated with reduced happiness and success. Indeed, learning-to-read an alphabetic script requires the explicit teaching of the correspondences between letters and sounds (phonics). Memorizing thousands of visual words by heart is no alternative as it would be like memorizing a telephone directory. Although most teachers indicate that they systematically teach letter-sound correspondences in grade 1, there is evidence that many of them complement explicit letter-sound teaching (strict phonics) with various whole word recognition strategies (mixed phonics). The goal of the present study was to find out whether strict phonics methods produce better results than less strict or mixed phonics methods. To do so, the responses of 9,340 grade 1 teachers concerning their teaching practice and the textbook they used were matched to the performance of their 139,288 students assessed in mid-grade 1 and early-grade 2 through the national evaluations in France. The results of hierarchical mixed effects modelling showed a clear advantage of strict phonics textbooks and strict phonics teaching methods over mixed, strongly mixed, or very strongly mixed (non-systematic) phonics methods while controlling for pre-reading skills, class reading level, socio-economic status, and teacher experience. The advantage of strict phonics methods over all other methods was even stronger for pupils with initially weak pre-reading skills and students from socio-economically disadvantaged neighborhoods. These results have far-reaching implications for educational policy, teacher training, prevention of reading difficulties, and social equality.

Introduction

Reading is the backbone of all learning. Poor reading is associated with low academic achievement, reduced professional success, low self-esteem, and sometimes more severe learning disabilities and psychiatric disorders [1–3]. Yet, even in industrialized countries with good access to education, around 20% of students encounter serious problems in learning to read and have difficulties understanding a written text when they enter high school [4]. This percentage is even higher for children from socially disadvantaged backgrounds [5–7]. Because all alphabetic writing systems transcribe spoken language, the initial stages of reading are all about cracking the alphabetic code, that is, the fact that letters correspond to the sounds of the spoken language, i.e., phonemes [8]. To master this initial stage of reading, high-quality explicit instruction is needed. Children cannot discover the rules that map the letters (graphemes) to the sounds of their language (phonemes) on their own. They need to be taught to do so [9, 10]. The effect of ineffective teaching cannot be easily compensated for by effective teachers in subsequent years, which underscores the importance of good teaching practice right from the start of schooling [11]. As argued by Reid Lyon and Weiser [12], “it is ineffective instruction that dooms children to a lifetime of reading failure” (p. 476).

Ever since the report of the National Reading Panel [13], there is a large consensus amongst reading scientists that the systematic teaching of grapheme-phoneme correspondences (i.e., systematic phonics approach) produces better results for all students than a non-systematic phonics approach or a non-phonics approach [14]. Although most teachers declare that they use a systematic phonics method,

they seem to do so with various levels of strictness, as suggested by the frequent use of global reading strategies (e.g., memorizing words). For example, in a questionnaire survey of about a thousand first-grade teachers in France [15], 91% of teachers stated that they "emphasize mastery of the correspondences between letters and sounds" but at the same time 82% of them declared that they asked their students to "recognize words holistically", and 85% stated that they taught student to "anticipate the meaning of words and anticipate words in a sentence" (p.66). A more recent survey of primary school teachers in Spain (N = 1716) showed that "most teachers favored whole-word methods... (and) showed inconsistent and limited knowledge of the evidence-based approach for effective reading instruction" [16, p.65]. More generally, there are serious doubts as to whether teaching practices are always in line with the recommendations of scientific research, whether teachers fully understand these components and whether they are able to implement them in their classrooms [17–21].

There is a hotly debated research question behind the issue of how strictly phonics methods should be conceived and applied. Should one exclusively teach grapheme-phoneme correspondences - a strict phonics approach - or should one add other reading strategies, such as whole-word memorization and contextual or semantic inference to guess the meaning of non-decodable words - a mixed phonics approach? [22] Proponents of a strict phonics approach warn of the risks involved in adding visual recognition or whole-word memorization strategies at the beginning of learning, which could lead pupils to develop "guess-reading" habits that might compromise the learning of a precise grapheme-phoneme correspondences. Proponents of a mixed phonics approach, on the other hand, argue that the introduction of a small number of "sight words" to be learned by rote, including frequent grammatical words with irregular pronunciation, pose no particular problem [23].

These questions are still widely debated today. The experimental studies published over the last twenty years do not currently allow us to draw any solid conclusions, mainly because most studies have simply not focused on the opposition between strict phonics and mixed phonics. In the most recent synthesis on the effectiveness of phonics methods, Torgerson, Brooks [24] concluded that all of the reviewed studies mixed code teaching and whole-word teaching to varying degrees. To our knowledge, the only study that compared a strict phonics method with a mixed phonics method that included memorizing frequent words by heart found no significant differences between the two methods [25]. However, this conclusion was weakened by the fact that the reading assessment at the end of the year included some of the words learned by rote in the mixed phonics program, which could have biased the results.

In the present study, we investigated whether strict phonics methods are better than mixed phonics methods in a mega-study of unprecedented scale. The strictness of the phonics method was operationalized by the textbooks teachers used and information about the implementation of the methods. For this purpose, a questionnaire was sent to all first-grade teachers of 17 French regional departments (N = 16,149) to find out about (1) teachers' characteristics (years of experience in the profession, level of education, etc.), (2) training courses attended by teachers during the three years preceding the survey, and (3) teaching practices in reading and writing, in particular, the textbook used, teaching of sight words, and the level of decodability of texts given to pupils to read. The teaching

methods were then classified into five groups on the basis of the strictness of the phonics approach, the number of sight words, and the level of decodability going from strict phonics to highly mixed phonics approaches (see below). The anonymized responses of the teachers who answered the survey (9,340, i.e., 1 in 5 teachers in France) were then matched with the anonymized results of the national evaluations [26] of their students in the middle of first grade and the beginning of second grade. This made it possible to investigate whether the strictness of phonics methods had an effect on students' subsequent performance in reading fluency and reading comprehension. Of particular interest was the question as to whether the strictness of the method mattered even more for children with weak pre-reading skills and children from socio-economically disadvantaged neighborhoods.

Results

Effect of Type of Textbook

The effects of the Type of Textbook on reading fluency and reading comprehension in mid-grade 1 (T1) and beginning grade 2 (T2) are presented in Figure 1. Note that we plot the estimated marginal means for an average school SES (5/10), good teacher experience (+ 5 years), average class reading level, and average individual pre-reading skills. The full statistical results of the mixed effect model with all standardized coefficients and p-values are found in Table 1.

As can be seen in Figure 1, students whose teachers use a strict phonics textbook had, on average, better scores than students learning with a mixed textbook of any other category (mixed, mixed+, mixed++) both on reading fluency and reading comprehension and at both assessment points (T1 and T2).

As can be seen in Table 1, at T1, learning to read with a strict phonics textbook resulted in an advantage between 7% to 12% of a standard deviation for reading fluency and 7% to 9% of a standard deviation for reading comprehension. As concerns T2, learning to read with a strict phonics textbook resulted in a 7% to 9% advantage for reading fluency and a 6% to 7% advantage in reading comprehension.

Table 1. Standardized regression coefficients and *p*-values resulting from hierarchical mixed effect models of Type of Textbook controlling for school SES, Teacher Experience, Class-reading level and Pre-reading skills.

	mid grade 1 (T1)				entry grade 2 (T2)			
	Reading Fluency		Reading Comprehension		Reading Fluency		Reading Comprehension	
	(1)	(2)	(3)	(4)	(3)	(4)	(3)	(4)
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
(Intercept)	-1.47***	<0.01	-1.71***	<0.01	-1.44***	<0.01	-1.55***	<0.01
school SES	0.02***	<0.01	0.03***	<0.01	0.05***	<0.01	0.05***	<0.01
Teacher Experience (< 5 ys)	<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>	
Teacher Experience (> 5 ys)	0.00	0.82	0.05***	<0.01	-0.01	0.267276	0.04***	<0.01
Class Reading Level	0.05***	<0.01	0.04***	<0.01	0.01**	0.03	-0.01	0.109
Pre-reading Skills	0.15***	<0.01	0.18***	<0.01	0.16***	<0.01	0.18***	<0.01
Strict Textbook	<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>	
Mixed Textbook	-0.12***	<0.01	-0.12***	<0.01	-0.08***	<0.01	-0.06**	<0.02
Mixed+ Textbook	-0.07**	<0.01	-0.08**	<0.01	-0.09***	<0.01	-0.06**	<0.02
Mixed++ Textbook	-0.10**	<0.01	-0.10***	<0.01	-0.07**	<0.01	-0.07**	<0.01
Number of Observations:	90 265		90 062		90 265		89 466	

Notes ***: p-value < 0.01; **: p-value < 0.05; *: p-value < 0.1

Note also that pre-reading skills at T0 had a strong effect on students' results in fluency and comprehension tests both at T1 and T2. A one-unit increase in our pre-reading variable (i.e., KAP) at T0 produced, on average, an increase of around 15% to 20% of a standard deviation in the fluency and comprehension tests at T1 and T2. The average class reading level and the school's SES had a small but significant effect: the higher the average class reading level or the socio-economic background of the school, the better the results of students at T1 and T2. Teacher's professional experience had a small positive effect but only on reading comprehension not on fluency.

Because the number of textbooks varied in each category, one could argue that there is a substantial internal heterogeneity within the different textbook categories, which could weaken the results obtained. We therefore carried out an additional analysis, for which we selected and introduced into the models only the most frequently used textbook from each category, enabling us to directly test the effect of these textbooks on students' results, while maintaining a sufficiently large number of classes. The results are equivalent to those presented above: pupils in classes using the strict phonics textbook achieved better results in reading than pupils in classes using mixed phonics textbooks (for detailed results, see SOM4, Table S2).

Effect of Type of Teaching Method

The Type of Teaching Method variable was based on the type of textbook used but also on the teaching practices declared by the teacher (see Methods). The effects of the Type of Teaching Method on reading fluency and reading comprehension in mid-grade 1 (T1) and grade 2 (T2) are presented in Figure 2. Note that we plot the estimated marginal means for an average school SES (5/10), good teacher experience (+ 5 years), average class reading level and average individual pre-reading skills. The full statistical results of the mixed effect model with all standardized coefficients and p-values are found in Table 2.

The results show that the strict phonics method was associated with significantly better reading fluency and comprehension in mid grade1 and early grade 2 than all other methods. As can be seen in Table 2, in mid-grade 1, learning to read with a strict phonics method was associated with a 11% to 16% of a standard deviation advantage in reading fluency and a 13% to 19% advantage in reading comprehension. In early grade 2, the advantage of the strict phonics method was between 9% and 11% of a standard deviation for reading fluency and 7% to 10% advantage for reading comprehension.

Table 2. Standardized regression coefficients and *p*-values resulting from hierarchical mixed effect modeling with Type of Teaching Method controlling for school SES, Teacher Experience, Class-reading level, and Pre-reading skills.

	mid grade 1 (T1)				entry grade 2 (T2)			
	Reading Fluency		Reading Comprehension		Reading Fluency		Reading Comprehension	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
(Intercept)	-1.43***	<0.01	-1.66***	<0.01	-1.43***	<0.01	-1.51***	<0.01
school SES T0	0.02***	<0.01	0.03***	<0.01	0.05***	<0.01	0.05***	<0.01
Teacher Experience (< 5 ys)	<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>	
Teacher Experience (> 5 ys)	0.00	0.87	0.05***	<0.01	-0.01	0.314	0.04***	<0.01
Class Reading Level T0	0.05***	<0.01	0.04***	<0.01	0.01**	0.028	-0.01	0.108
Pre-reading Skills T0	0.15***	<0.01	0.18***	<0.01	0.16***	<0.01	0.18***	<0.01
Strict Method	<i>ref</i>		<i>ref</i>		<i>ref</i>		<i>ref</i>	
Mixed Method	-0.16***	<0.01	-0.19***	<0.01	-0.11**	0.013	-0.10**	0.013
Mixed+ Method	-0.11**	0.009	-0.13**	<0.01	-0.11***	<0.01	-0.08**	0.020
Mixed++ Method	-0.15***	<0.01	-0.15**	<0.01	-0.09*	0.054	-0.07*	0.085
Other	-0.12***	<0.01	-0.14***	<0.01	-0.10***	<0.01	-0.10***	<0.01
Number of Observations:	90 265		90 062		90 265		89 466	

Notes ***: *p*-value < 0.01; **: *p*-value < 0.05; *: *p*-value < 0.1

Are Strict Phonics Methods More Efficient For Initially Weak Readers?

Previous studies suggested that systematic phonics approaches are particularly beneficial for readers with initially weak reading skills [14, 27, 28]. This possibility was addressed in our analyses by introducing an interaction term between Type of Teaching Methods and Pre-reading Skills at T0, which were measured by KAP at the individual level. The interaction between the effects of Pre-reading Skills (KAP) and Type of Method was highly significant in all cases [Fluency T1: $\chi^2(4) = 26.776$, $p < .01$; RC T1: $\chi^2(4) = 22.126$, $p < .01$; Fluency T2: $\chi^2(4) = 28.451$, $p < .01$; RC T2: $\chi^2(4) = 14.786$, $p < .01$]. The results are plotted in Figure 3 and full statistical analyses are provided in Table S3 in the online supporting materials.

As can be seen in Figure 3, the weaker the initial individual pre-reading level, the greater the advantage of the strict phonics method over the other methods. For example, for pupils with a high level of KAP at the start of grade 1 (8/10), switching from a strict phonics method to a mixed method reduces fluency levels by 10.3% of a standard deviation at T1 and 10.1% of a standard deviation at T2. For students with the

same characteristics but a low level of KAP (2/10), switching from a strict phonics method to a mixed method reduces fluency levels by 14.7% of a standard deviation at T1 and 14.7% of a standard deviation at T2. The same pattern is found for reading comprehension. For pupils with a high level of KAP at the start of grade 1 (T0), the difference in effectiveness between methods is estimated at 11.4% of a standard deviation at T1 and 6.7% at T2; for those with a low level of KAP, the differences between methods increase to 20.5% of a standard deviation at T1 and 14.7% of a standard deviation at T2.

Are Strict Phonics Methods More Efficient For Beginning Readers from Social-Disadvantaged Backgrounds?

Previous research has shown that strict phonics methods are particularly efficient for readers from socio-economically disadvantaged neighborhoods [14, 27]. To assess this potential effect in our sample, we added the interaction between Type of Teaching Method and SES to the mixed-effects model. The interaction between the effects of Type of Teaching Method and SES was indeed statistically significant for all outcome measures (all $ps < .01$, see SOM 6), thus indicating that the effects of the Type of Teaching Method vary as a function of SES with a greater advantage of the strict phonics method for low SES classrooms (see Table S4). For example, in schools with relatively high school SES (category 7/10), the gap in reading fluency between the strict method and the mixed method is estimated at 8% of a standard deviation at T1 and at 4% of a standard deviation at T2. The difference between these two methods is much greater in schools with a low SES (category 2/10), with 17% of a standard deviation at T1 and 23% at T2. The same is true for reading comprehension: the advantage of the strict method over the other methods is always greater in low SES than in high SES classes both at T1 and T2.

Discussion

This is the first study to clearly demonstrate that the type of textbook and the type of teaching method used during the first months of grade 1 reading instruction significantly affects reading fluency and reading comprehension later on. Specifically, strict phonics textbooks and strict phonics teaching methods produce superior levels of reading fluency and reading comprehension than less strict, mixed, or highly mixed phonics methods. The advantage of the strict phonics method over the other methods was greater for children who were initially weaker readers at the start of grade 1 and who came from socioeconomically disadvantaged neighborhoods. The effects of type of textbook and type of teaching method were obtained controlling for other variables, such as teachers' experience, initial pre-reading level, class reading level, or SES. The longitudinal nature of the study allows us to draw causal conclusions about the learning progress of children as a function of type of textbook and the type of teaching method. The scale of the study with more than 6,400 teachers and their 90,000 pupils is unprecedented making the results extremely robust and representative. The results have far-reaching implications about efficient teaching of reading in the classroom not only in France but also in other countries, as all writing systems offer the possibility to mix strict phonics methods with other methods.

Our results provide very strong support for the conclusions of the NRP suggesting that the systematic teaching of grapheme-phoneme correspondences (systematic phonics approach) produces better results for all students than a non-systematic phonics approach or a non-phonics approach [14]. Although there is a large consensus about the recommendations of the NRP in the scientific community, the results of our survey join those of Suárez, Sánchez [21] to suggest that teachers not always use practices based on the recommendations of the NRP, as many of them declare using whole-word reading strategies as much as grapheme-phoneme teaching.

Our study provides further evidence on a point that goes beyond the scientific consensus, that is, is it important and beneficial to introduce a number of non-decodable irregular words right from the start? Some authors have indeed suggested that it does not matter to introduce irregular words as long as teaching is based on introducing grapheme-phoneme correspondences in a systematic way [23]. However, our results clearly suggest that this is not true. Having children memorize irregular words, giving them not fully decodable texts to read, or spending time on global recognition of non-decodable words are associated with weaker reading fluency and comprehension performance on year later, and even more so for children with weak pre-reading skills or from socio-economically disadvantaged neighborhoods.

Why Is Strict Phonics Better Than Mixed Phonics?

According to one of the traditional models of learning-to-read, Frith's [29] three-stage model, beginning readers go through three stages on their way to becoming fluent readers, the logographic stage, the alphabetic stage and the orthographic stage. In the logographic stage, the words are processed just like any other visual object or symbol and their meanings are associated with global visual shapes and features. In the alphabetic stage, children learn to associate graphemes to phonemes in order to access words in their phonological lexicons (i.e., decoding). In the final orthographic stage, repeated access of the same words allows the child to create orthographic representations of words, which form the basis for fast, parallel and automatic word recognition. According to this model, mixed phonics methods might fare less well because they might prevent children from moving from the (inefficient) logographic stage to the (productive) alphabetic stage or they might make children believe that they could skip the alphabetic stage and move right onto the orthographic stage. However, the very existence of a logographic stage has been questioned [30] and doubts have been raised as to whether the model is a good model of learning to read across languages and writing systems [31].

The present results can also be interpreted in the context of state-of-the-art models of learning to read alphabetic scripts. One of the most powerful models is based on Share's [32] phonological decoding and self-teaching theory. The core idea is that phonological decoding provides a powerful self-teaching device because the explicit learning of a small set of spelling-sound correspondences allows the child to decode an increasingly large number of words. The theory is consistent with the idea that phonological decoding is at the heart of reading acquisition across all alphabetic writing systems [8, 33]. The more transparent the mapping is between graphemes and phonemes, the easier it can be learnt and taught.

Indeed, the transparency of the writing system perfectly predicts the rate of reading acquisition across alphabetic writing systems at the end of grade 1 [34, 35]. Share's phonological decoding and self-teaching theory has been implemented in a computational model [36] that learns just like a child through an explicit teaching phase, where the model learns to associate graphemes with phonemes (i.e., decoding network) followed by unsupervised self-teaching. Given that the decoding network is a (linear) two-layer associative network that cannot handle irregular words (i.e., no hidden units), the presentation of irregular words early on produces "catastrophic interference" in the decoding network [37], that is, interference with learning stable grapheme-phoneme associations. Therefore, strict phonics is superior for the initial configuration of the decoding network, which will bootstrap subsequent orthographic learning [36].

Interestingly, a recent analysis of the exhaustive French national evaluations [38] also showed that a variable associated with the rudimentary decoding network, knowledge of the alphabetic principle (KAP), was a more powerful predictor of subsequent reading fluency and comprehension than either letter knowledge or phoneme awareness, which underscores the importance of the mapping process (i.e., the teaching method) over and above the quality of input or output representations (i.e., letters, phonemes). The interaction we found between KAP and Type of Teaching method goes in the same direction. That is, we showed that children with high levels of KAP were less affected by mixed methods than those with low levels of KAP, presumably because their decoding network is already well in place and can resist to "catastrophic interference". However, this is not the case for children with weak levels of KAP, who will be much more affected by using less strict or mixed phonics methods.

The Effect of Social Environment

Our results indicate that the effect of teaching method also varies according to the social context of the schools. A strict phonics method is even more effective in schools with low SES than in schools with high SES. School SES is highly correlated with the parents' level of education. Given that the statistical modelling controlled for both the pre-reading level at entry and the average class reading level, which has the effect of capturing a large part of the effect of both the school and the social context of the schools, it is likely that the SES effect reflects parents' level of education and educational support. It is well known that the number of hours parents spend reading with their children and the number of books in the home are positively associated with SES [39] and children from higher SES backgrounds are more likely to own books and to have greater access to educational resources [40]. Thus, SES is clearly associated with home literacy environment, which is related to the parents' level of education. We suggest that, in privileged social environments, the difference between teaching methods is reduced, since parental intervention tends to compensate the level of students who learn in classes using less effective teaching methods. Conversely, the gap between teaching methods is greatest in socially disadvantaged contexts, where parental intervention is less frequent or less effective [41].

Conclusion

The present study clearly demonstrated that the type of textbook and teaching method used in the first few months of primary school has a definite effect on subsequent reading fluency and reading comprehension, especially for those who enter school with weak pre-reading skills or who come from a low socio-economic background. Yet, the strict phonics method, which turned out to be the most effective in the present study, is currently very rarely used in French classrooms. This is rather alarming, and we suspect that this is not very different in many other countries (see the recent survey by Giménez, Sánchez [16] on teaching practices in Spain).

A good teaching method is essential to compensate for social inequalities that are omnipresent in many countries. In our study, pupils from socially-disadvantaged schools (SES = 3) who receive strict synthetic teaching do, in fact, achieve better results in the mid-grade fluency test than pupils from socially advantaged schools (SES = 7) who learn to read using a mixed method, all other variables being equal. An effective policy for combating inequalities in early learning thus needs to put an emphasis on improving the effectiveness of teaching.

The success of such an undertaking requires resolute action and substantial resources, at all relevant levels. The first level to consider is that of official instructions. While school curricula for the first year of primary school rightly insist on the mastery of grapheme-phoneme correspondences, they would probably benefit from specifying the means to achieve them. The second level of action concerns teaching supports. As shown above, the most effective textbooks are only used by a minority of teachers in France. There is little doubt that a significant improvement in the quality of available textbooks would enable teachers to increase the effectiveness of their teaching activities. A third level of action, which we believe to be decisive, lies in the training of teachers and educational advisors. Today, the professional culture of reading instruction is still marked by old pedagogical principles whose ineffectiveness has been demonstrated by research. This is not unique to France, as it is the case in many comparable countries [16, 21, 23, 42]. Only a substantial investment in initial and in-service teacher training could change the situation. The results of our survey show that the pedagogical quality of the textbook is important, but so is the way it is used. Enhanced pedagogical support, informed by research findings, is absolutely essential to help teachers put in place effective professional gestures in the teaching of reading.

In sum, clarifying the expectations of official programs, enabling teachers to choose high-quality textbooks, developing training programs that are closer to the realities of the classroom are undoubtedly the levers that would give teachers the means to exercise informed pedagogical freedom when it comes to teaching reading. The stakes are high, not only in terms of improving the results of the education system, but also in terms of social justice because far too many pupils, often from low socio-economic backgrounds, are “doomed to a life of reading failure” [12] already at the end of the first year of primary school.

Methods

Sample

The data collection phase began in January 2021. All first-grade teachers in public and private schools under contract in 17 regional departments of France were invited to respond to an online questionnaire (see below) sent by the *École Normale Supérieure* to all schools in these regional departments. The 16,149 first-grade teachers in the departments concerned were contacted for the first time between January 19 and 26, 2021, resulting in 4,357 responses to the questionnaire. An initial follow-up survey was carried out the following week, resulting in 8,821 complete responses. Following two further follow-up phases during the first two weeks of February, 9,340 teachers finally completed the questionnaire in full, giving a final response rate of 59%, representing around one in five grade 1 teachers in France. The average time taken to complete the questionnaire was 11 minutes. The anonymized responses of the 9,340 teachers surveyed were then matched by the *Department for evaluation, foresight and performance* (DEPP), which belongs to the French Ministry of Education, in January 2022 to the equally anonymized results of the national assessments [26] of their 139,288 pupils enrolled in 1st grade during 2020-2021 at three assessment points (grade 1 entry, grade 1 middle and grade 2 entry).

For the present analyses, we excluded teachers who declared working part time (less than 75%), thus ensuring that the teaching methods described by teachers in the questionnaire were indeed those to which their pupils had been exposed in class during the school year. This resulted in the exclusion of 762 respondents. We further excluded those who declared that they did not use any textbook to teach reading, corresponding to 2,138 teachers. The present analyses are thus based on a final sample of the 6,440 teachers as well as their 90,265 pupils. The final sample is representative of all first-grade classes in France in terms of the social position index of the schools, the average class level of knowledge of the alphabetic principle (KAP), and the average student level of knowledge of the alphabetic principle (see Table S1 for details).

Questionnaire

The development of the online questionnaire was based on the results of classroom observations carried out during the two academic years 2018-2019 and 2019-2020 [43], which served as the basis for formulating the questions addressed to the teachers. The survey inquired about the characteristics of teachers and the reading practices they employed during the initial three months of the academic year. A description of the questions is presented in the supporting on-line materials (SOM2).

Type of textbooks

The 34 most frequently used textbooks (i.e., 95% across all classes that used textbooks) were submitted to a systematic analysis of their characteristics on the basis of two criteria:

(1) The decodability rate of the texts given to pupils as calculated by the *Anagraph* program[1] (for details, see SOM3).

(2) The number of non-decodable words to be learned overall by pupils over the same period.

The results for all 34 textbooks along with their frequency of usage can be found in Appendix A. As can be seen in Appendix A, the decodability rate of texts given to read varied enormously from one textbook to another, ranging from 44% to 100%. The same applied to the number of words given to memorize overall at the start of the year, which varies from 0 to 50 words. The correlation between these two characteristics was relatively high ($r = - 0.54$): the less decodable the textbook, the greater the number of words to be memorized globally. Based on these two criteria, the 34 textbooks were classified into four categories:

STRICT: strict phonics textbooks (5% of classes). This first category includes textbooks with a text decodability rate strictly above 95%, and no non-decodable words to be learned globally.

MIXED: mixed phonics textbooks (9% of classes). This second category includes textbooks with text decodability equivalent to that of the first category, but also with non-decodable words to be learned globally (10 on average).

MIXED+: strongly mixed textbooks (74% of classes). This third category includes textbooks with a decodability between 76% and 95%, and non-decodable words to be learned globally (21 on average).

MIXED++: very strongly mixed phonics textbooks (12% of classes). This fourth category includes textbooks that yield text decodability below 76% and non-decodable words to be learned globally (28 on average). This category can be considered as nonsystematic phonics.

Type of Teaching Method

Studying the effect of the type of textbook is a first way of approaching the effect of teaching methods for learning-to-read. However, it is possible that teachers use practices that deviate from the textbook's recommendations. That is why we added a second variable to better characterize the teaching method actually used in the classroom. The "type of method" variable was thus constructed on the basis of the type of textbook used, the degree to which the teachers declared that they followed the recommendations in the textbook and, finally, the declared teaching practices concerning the decodability of texts given to read in class and the number of non-decodable words to be learnt by heart. This resulted in five categories for the "type of method" variable:

STRICT (n = 96 classes): Teachers declared that they use a strict phonics textbook (category 1), strictly followed its recommendations, did not make their pupils memorize non-decodable words, and gave their pupils exclusively fully decodable texts to read.

MIXED: (n = 187 classes): Teachers declared that they used a mixed textbook (category 2), strictly followed its recommendations, had their pupils memorize non-decodable words to a certain extent and read texts that were not fully decodable.

MIXED+ (n = 2,016 classes): Teachers declared that they used a strongly mixed textbook (category 3), strictly followed its recommendations, had their pupils memorize non-decodable words and read texts

that were not fully decodable.

MIXED++: (n = 194 classes): Teachers declared that they used a very strongly mixed textbook (category 4), strictly followed its recommendations, had their pupils memorize non-decodable words and read texts that were not fully decodable.

Other (n = 3,947 classes): This category included teachers who did not follow the recommendations of their textbook, whether it was strict or mixed and/or declared practices that contradict the principles of the textbook they used.

Reading and Control Variables

Reading Fluency (T1, T2)

Reading fluency is one of the main outcome variables of the exhaustive national evaluations [26] measured in mid-grade 1 (T1) and at the beginning of grade 2 (T2). The fluency tests were administered individually by teachers to their students, who were asked to read a text for one minute. Teachers reported the number of correctly read words. The mid-grade 1 text had 30 words, while the grade 2 text had 103 words. The fluency variables used in the models therefore correspond to the number of words correctly read in one minute.

Reading Comprehension (T1, T2)

Two subtests were used to assess reading comprehension (RC) in the national evaluations [26]. In mid-grade 1 (T1), pupils were asked to read sentences and then circle the picture corresponding to their meaning. This test contained eight items and, for each item, four pictures were proposed.

In grade 2 (T2), the same RC test as above was administered again using 10 different sentences. In addition, a second test was added, in which students were asked to read two texts and answer four multiple-choice questions at the end of the text. Thus, the RC variable at T2 correspond to the total number of correct answers obtained by the students in the two subtests (i.e., a maximum of 18 correct answers).

Pre-Reading Skills At The Start Of First Grade (T0)

Pre-reading skills at the start of first grade are an important determinant of the reading level achieved at the end of the year. The national evaluations at the beginning of grade 1 (T0) contain several measures: knowledge of the letters of the alphabet, oral comprehension, phonemic awareness and, finally, knowledge of the alphabetic principle (KAP). We have chosen KAP as a measure for pre-reading skills because we have previously shown that this variable explains reading fluency one year later better than all other pre-reading or reading-related measures [38]. KAP is measured by a 10-item grapho-phonemic linking exercise. For each item, the teacher pronounces a word aloud, and pupils are asked to circle the one out of five letters (e.g., m, n, b, d, f) that corresponds to the phoneme with which the word begins. For

example, if the word in the test item is "feuille" (leaf), the series of letters proposed to the students is "m, n, b, d, f" and the correct response is "f".

Teacher's Years of Experience in Grade 1 (T0)

It is a well-established finding that teacher effectiveness varies as a function of their professional experience. However, it seems that apart from the first few years in the profession, additional years of experience seem to have no significant effect on students' performance [44, 45]. Thus, the number of years of professional experience provided by the teachers via the questionnaire was transformed into a dichotomous variable that distinguishes novice (1 to 4 years of classroom experience) from more expert teachers (more than 4 years of classroom experience).

Class Reading Level at Entry of Grade 1 (T0)

Because numerous studies have shown that teachers tend to lower their standards when faced with weak classes [46], we included the average reading level of each class to control for this potential effect that might affect teaching practice. Given that the best pre-reading variable was KAP [38], the average class reading level was constructed from the average KAP of the pupils in each class.

School Socio-economic Status (T0)

The second variable used to characterize the context of each class was the school's socio-economic status (SES). This index, calculated by the *DEPP*, characterizes the social composition of all schools. The greater the number of privileged families in a given school, the higher the value of the school SES. We constructed a 10-level variable according to the SES deciles of the schools in our sample. The first SES category corresponded to schools in the first SES decile - i.e. the 10% of schools with the most disadvantaged background; the second category groups schools between the first and second decile, and so on. For example, in the first SES category, the proportion of families with at least one parent with a higher education qualification is 18%. In the fourth SES category, this proportion rises to 38%, and in the seventh to 56%. While the school SES cannot fully replace a direct measurement of each student's social and cultural background, it is nonetheless a very good proxy for it.

Statistical analyses: multi-level modeling

Our data are structured on two levels. The first level groups together variables concerning the pupils themselves, such as pre-reading level at T0. The second level groups variables that concern the class, that is, variables that take the same value for all pupils in a given class. This is the case, for example, for the average class reading level, years of teacher experience, or SES. This is of course also the case for the type of textbook used or the teaching method implemented in the classroom, which are our variables of interest here.

In the presence of this type of data, it is customary to implement multilevel modeling, which allows us to take into account the fact that the data have a hierarchical structure when estimating the model's

coefficients [47]. Regression coefficients were estimated using multilevel linear random-effects modeling, which introduces structured variables into the model at both individual (student) and group (class) levels. The general form of the multi-level regression model implemented is therefore as follows:

1. Individual Reading Ability = School SES + Teacher Experience + Average class reading level + Individual Pre-reading level + Type of textbook or Type of method

The estimation of the parameters was done using a multilevel linear random-effects model with students as level 1 and class as level 2. The coefficients were estimated using the lme4 package version 1.1-32 of the R software [48].

[1] <https://anagraph.ens-lyon.fr/app.php>

Declarations

Statement of consent. The school teachers who participated in the on-line survey gave informed consent prior to participating in the survey. The reading performance of their pupils was extracted post-hoc from the French national evaluations conducted by the French Ministry of Education. The national evaluations are exhaustive and obligatory for all students in France. The Ministry is not obliged to obtain students' consent. Access to the anonymized database that contains the results of the national evaluations and respect of general data protection rules (GDPR) was governed by a joint research agreement (Convention # 2020-13) between the researchers and the Department of Evaluation, Foresight and Performance (DEEP) of the French Ministry of National Education in charge of the national evaluations.

References

1. Smart, D., et al., *Consequences of childhood reading difficulties and behaviour problems for educational achievement and employment in early adulthood*. British Journal of Educational Psychology, 2017. **87**(2): p. 288-308.
2. Maughan, B., et al., *Reading Problems and Depressed Mood*. Journal of Abnormal Child Psychology, 2003. **31**(2): p. 219-229.
3. Francis, D.A., et al., *The association between poor reading and internalising problems: A systematic review and meta-analysis*. Clinical Psychology Review, 2019. **67**: p. 45-60.
4. OECD, *PISA 2018 Results (Volume I)*. 2019.
5. Billard, C., et al., *Reading, spelling and comprehension level in low socioeconomic backgrounds: Outcome and predictive factors*. Revue D Epidemiologie Et De Sante Publique, 2010. **58**(2): p. 101-110.
6. Fluss, J., et al., *Poor Reading in French Elementary School: The Interplay of Cognitive, Behavioral, and Socioeconomic Factors*. Journal of Developmental and Behavioral Pediatrics, 2009. **30**(3): p. 206-216.

7. Noble, K.G., M.J. Farah, and B.D. McCandliss, *Socioeconomic background modulates cognition-achievement relationships in reading*. Cognitive Development, 2006. **21**(3): p. 349-368.
8. Ziegler, J.C. and U. Goswami, *Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory*. Psychological Bulletin, 2005. **131**(1): p. 3-29.
9. Foorman, B.R., et al., *The role of instruction in learning to read: Preventing reading failure in at-risk children*. Journal of Educational Psychology, 1998. **90**(1): p. 37-55.
10. Rupley, W.H., T.R. Blair, and W.D. Nichols, *Effective Reading Instruction for Struggling Readers: The Role of Direct/Explicit Teaching*. Reading & Writing Quarterly, 2009. **25**(2-3): p. 125-138.
11. Sanders, W.L. and J.C. Rivers, *Cumulative and residual effects of teachers on future student academic achievement*. 1996: Knoxville: University of Tennessee.
12. Reid Lyon, G. and B. Weiser, *Teacher Knowledge, Instructional Expertise, and the Development of Reading Proficiency*. Journal of Learning Disabilities, 2009. **42**(5): p. 475-480.
13. NRP, *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction: Reports of the Subgroups*. 2000, National Institute of Child Health and Human Development.: Bethesda, MD.
14. Ehri, L.C., et al., *Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis*. Review of Educational Research, 2001. **71**(3): p. 393-447.
15. Fijalkow, É. and J. Fijalkow, *Enseigner à lire-écrire au CP : état des lieux*. Revue française de pédagogie, 1994(107): p. 63-79.
16. Giménez, A., et al., *Teachers' Opinions about the Teaching of Reading in Spain*. Psicología Educativa, 2023. **29**: p. 65 - 73.
17. Cunningham, A.E., et al., *How Teachers Would Spend Their Time Teaching Language Arts: The Mismatch Between Self-Reported and Best Practices*. Journal of Learning Disabilities, 2009. **42**(5): p. 418-430.
18. Joshi, R., et al., *Do Textbooks Used in University Reading Education Courses Conform to the Instructional Recommendations of the National Reading Panel?* Journal of Learning Disabilities, 2009. **42**(5): p. 458-463.
19. Podhajski, B., et al., *Professional Development in Scientifically Based Reading Instruction: Teacher Knowledge and Reading Outcomes*. Journal of Learning Disabilities, 2009. **42**(5): p. 403-417.
20. Kaiser, L., S. Rosenfield, and T. Gravois, *Teachers' Perception of Satisfaction, Skill Development, and Skill Application After Instructional Consultation Services*. Journal of Learning Disabilities, 2009. **42**(5): p. 444-457.
21. Suárez, N., et al., *Is Reading Instruction Evidence-Based? Analyzing Teaching Practices Using T-Patterns*. Frontiers in Psychology, 2018. **9**.
22. Landerl, K., *Influences of orthographic consistency and reading instruction on the development of nonword reading skills*. European Journal of Psychology of Education, 2000. **15**: p. 239-257.

23. Castles, A., K. Rastle, and K. Nation, *Ending the Reading Wars: Reading Acquisition From Novice to Expert*. Psychological Science in the Public Interest, 2018. **19**(1): p. 5-51.
24. Torgerson, C., et al., *Phonics: reading policy and the evidence of effectiveness from a systematic 'tertiary' review*. Research Papers in Education, 2019. **34**(2): p. 208-238.
25. Shapiro, L.R. and J. Solity, *Differing effects of two synthetic phonics programmes on early reading development*. British Journal of Educational Psychology, 2016. **86**(2): p. 182-203.
26. Andreu, S., et al., *Évaluations Repères 2022 de début de CP et de CE1*, in *Note d'Information 2023*, DEPP. p. 1-4.
27. Machin, S., S. McNally, and M. Viarengo, *Changing How Literacy Is Taught: Evidence on Synthetic Phonics*. American Economic Journal: Economic Policy, 2018. **10**(2): p. 217-41.
28. McArthur, G., et al., *Phonics training for English-speaking poor readers*. Cochrane Database of Systematic Reviews, 2018(11).
29. Frith, U., *Beneath the surface of developmental dyslexia*, in *Surface Dyslexia*, K.E. Patterson, J.C. Marshall, and M. Coltheart, Editors. 1985, Lawrence Erlbaum. p. 301-330.
30. Wimmer, H. and P. Hummer, *How German-speaking first graders read and spell: Doubts on the importance of the logographic stage*. Applied Psycholinguistics, 1990. **11**(4): p. 349-368.
31. Sprenger-Charolles, L. and P. Bonnet, *New doubts on the importance of the logographic stage*. Current Psychology of Cognition, 1996. **15**: p. 173-208.
32. Share, D.L., *Phonological recoding and self-teaching: Sine qua non of reading acquisition*. Cognition, 1995. **55**(2): p. 151-218.
33. Ziegler, J.C. and U. Goswami, *Becoming literate in different languages: similar problems, different solutions*. Developmental Science, 2006. **9**(5): p. 429-36.
34. Seymour, P.H.K., M. Aro, and J.M. Erskine, *Foundation literacy acquisition in European orthographies*. British Journal of Psychology, 2003. **94**: p. 143-174.
35. Ziegler, J.C., C. Perry, and M. Zorzi, *Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models*. Current Directions in Psychological Science, 2020. **29**(3): p. 293-300.
36. Ziegler, J.C., C. Perry, and M. Zorzi, *Modelling reading development through phonological decoding and self-teaching: implications for dyslexia*. Philosophical Transactions of the Royal Society B: Biological Sciences, 2014. **369**(1634).
37. McCloskey, M. and N.J. Cohen, *Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem*. Psychology of Learning and Motivation, 1989. **24**: p. 109-165.
38. Gioia, P., J.C. Ziegler, and J. Deauvieau, *Revisiting the causal effects of phonemic awareness on reading acquisition: insights from a systematic review and a large-scale longitudinal study*. PsyArXiv, 2024.
39. Raz, I.S. and P. Bryant, *Social background, phonological awareness and children's reading*. British Journal of Developmental Psychology, 1990. **8**(3): p. 209-225.

40. Bradley, R.H., et al., *The Home Environments of Children in the United States Part I: Variations by Age, Ethnicity, and Poverty Status*. Child Development, 2001. **72**(6): p. 1844-1867.
41. Morais, J., *Lire, écrire et être libre: de l'alphabétisation à la démocratie*. 2015: Odile Jacob.
42. Kim, J.S., *Research and the Reading Wars*. Phi Delta Kappan, 2008. **89**(5): p. 372-375.
43. Gioia, P., *L'entrée dans l'écrit. Etude croisée des pratiques scolaires et d'éducation familiale*, in *Sociology*. 2024, Ecole normale supérieure - Paris Sorbonne Lettres: Paris.
44. Mueller, S., *Teacher experience and the class size effect – Experimental evidence*. Journal of Public Economics, 2013. **98**: p. 44-52.
45. Chetty, R., J.N. Friedman, and J.E. Rockoff, *Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates*. American Economic Review, 2014. **104**(9): p. 2593–2632.
46. Terrail, J.-P., *De l'inégalité scolaire*. 2002, Paris: Éditions la Dispute.
47. Raudenbush, S.W. and A.S. Bryk, *Hierarchical linear models: Applications and data analysis methods*. Vol. 1. 2002: Sage.
48. Bates, D., et al., *Fitting Linear Mixed-Effects Models Using lme4*. Journal of Statistical Software, 2015. **67**(1): p. 1 - 48.

Figures

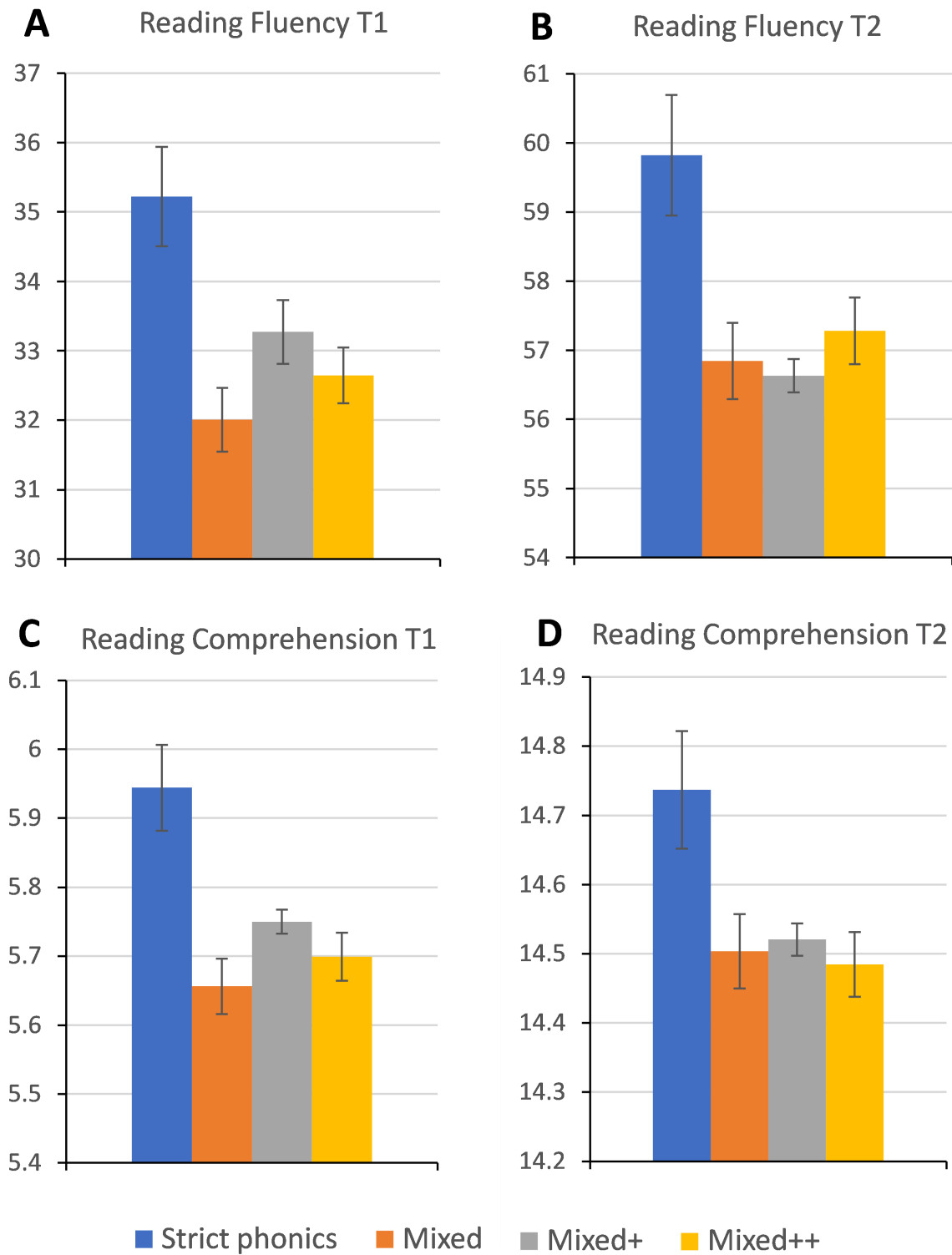


Figure 1

Estimated marginal means for the effect of textbook on reading fluency (words per minute) at T1 and T2 (A, B) and reading comprehension (number of correctly responded items) at T1 and T2 (C, D). Error bars represent standard errors.

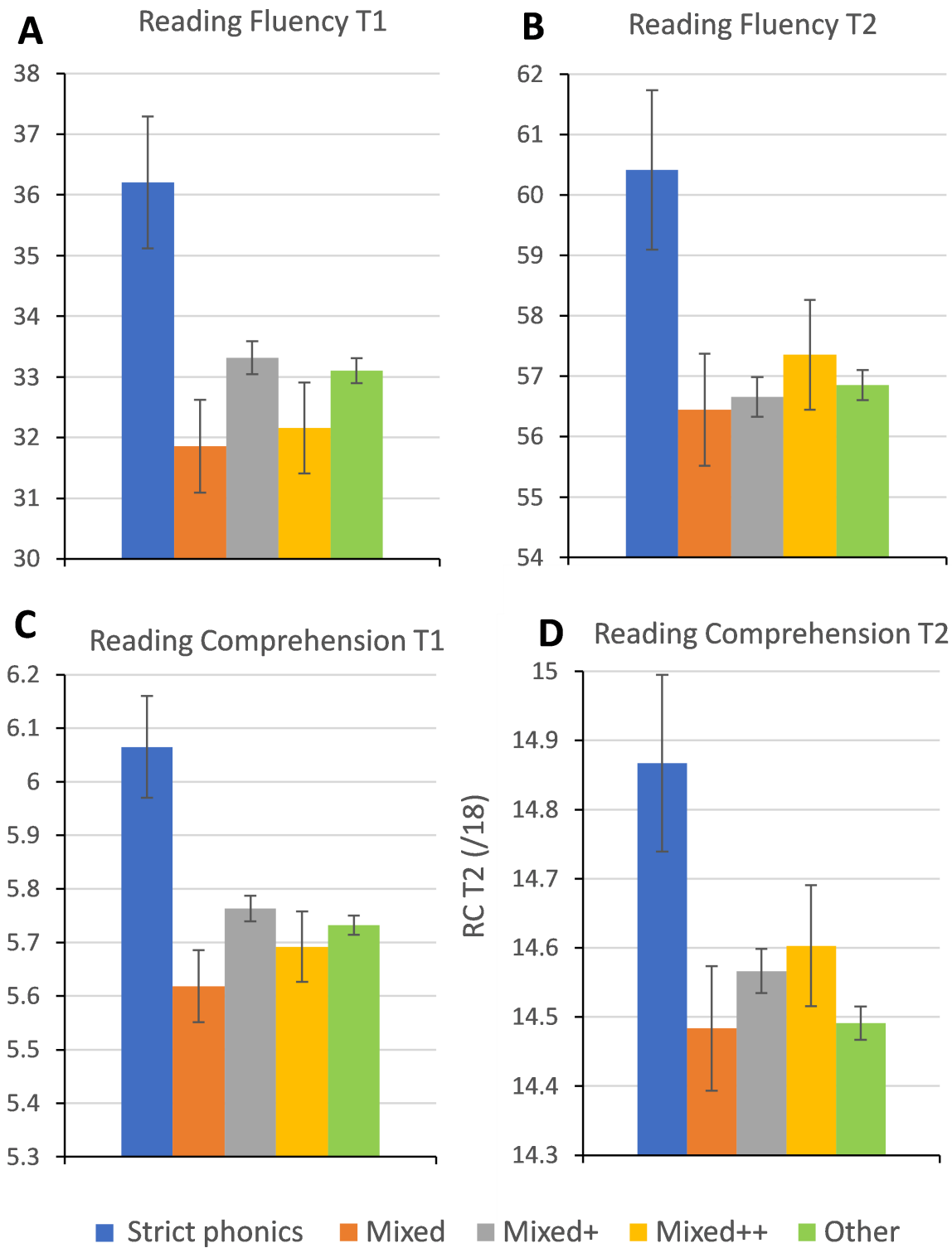


Figure 2

Estimated marginal means for the effect of Teaching Methods on reading fluency at T1 and T2 (A, B) and reading comprehension at T1 and T2 (C, D). Error bars represent standard errors.

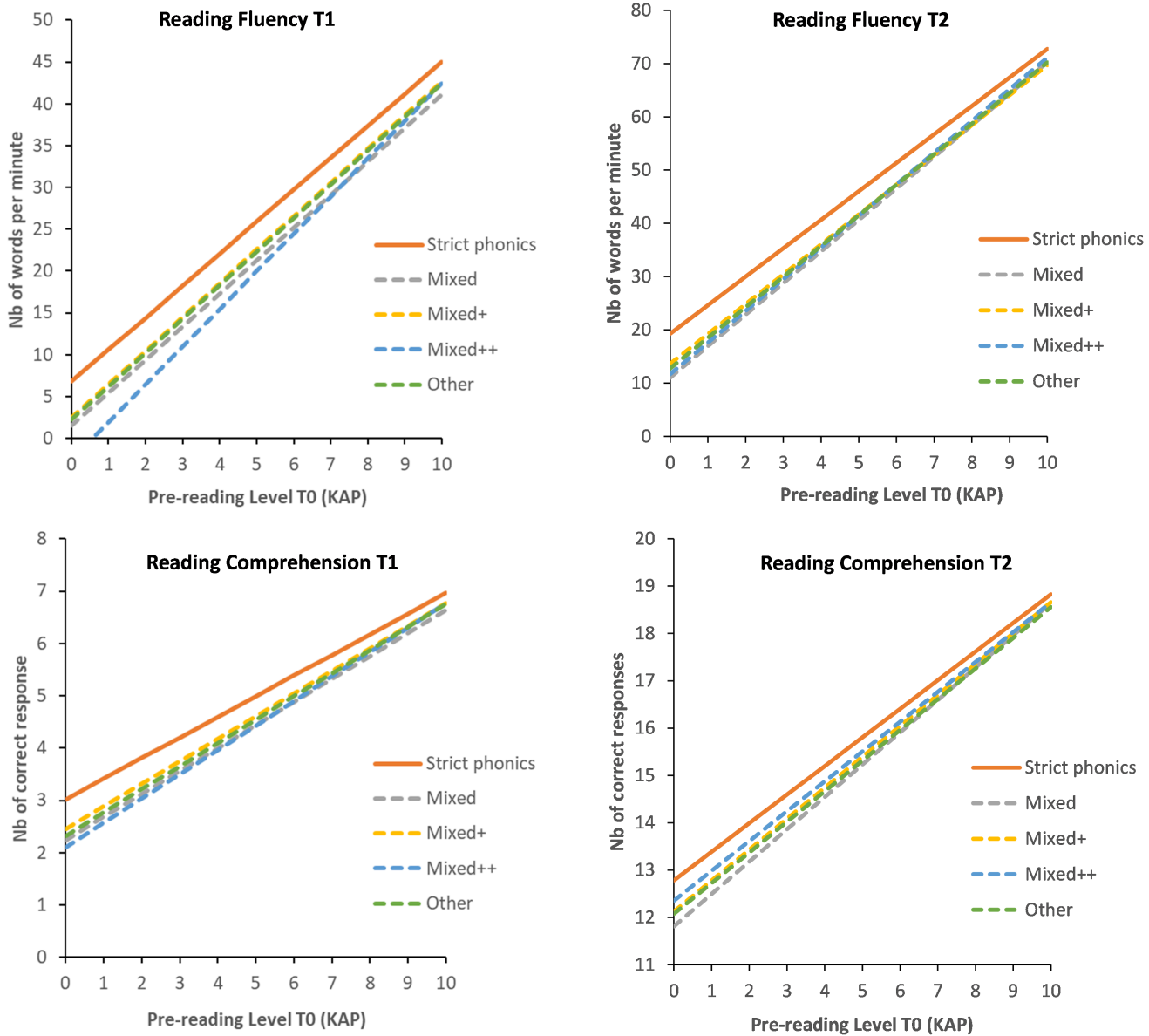


Figure 3

Interaction plots between the effects Pre-reading level at T0 (KAP) and Type of Teaching Method on Reading Fluency and Reading Comprehension in mid-grade 1 (T1) and entry grade 2 (T2).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupportingonlinematerialsforNAtureCommfinal.docx](#)
- [nrreportingsummary.pdf](#)
- [AppendixA.docx](#)