



Beginning reading and writing

Roland Goigoux

► To cite this version:

Roland Goigoux. Beginning reading and writing: Study on the influence of teaching practices on quality of basic learning . 2017. hal-01487491v1

HAL Id: hal-01487491

<https://hal.science/hal-01487491v1>

Preprint submitted on 12 Mar 2017 (v1), last revised 28 Aug 2017 (v2)

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.



Institut Français de l'Éducation, École Normale Supérieure de Lyon, France



Beginning reading and writing

Study on the influence of teaching practices on quality of basic learning

Synthesis of the research report

Directed and supervised by Roland Goigoux

Professor at Université Clermont Auvergne



A. Focal issue and methodology

A.1 Objectives, context and genesis of *Lire-Écrire au CP*—the national project for first-grade Teaching to Read-and-Write policy

Since the early 2000s, first-world governments have increasingly been turning to scientific research to ground their education policy on reliable empirical data (or ‘evidence-based education’, Davies, 1999; Saussez & Lessard, 2009). In France, policy-making on initial teaching to read and write is centrally concerned with finding the right methods—the “time-tested” methods (Cusset, 2014; Rey, 2014). The aim is to identify the didactic strategies that work best, to round off, in this specific domain, the wider evidence already available on the instructional strategies that work best for teachers’ effectiveness (Bressoux, 2007; Hattie, 2009, 2012).

After a long history of leaning on their education ministry authorities to define “best practice” (IGEN, 1995, 2006), successive French governments turned to the research community in a move to scaffold educational programs in strength of evidence and escape the draw of faddism and fashion (Slavin, 2008; Blanquer, 2014). This shift in direction came in the wake of Antoine Prost’s advocacy report on education research (2001) which made it plain that the spiralling number of qualitative studies on a small number of classes did not make up for the lack of studies on large cohorts of teachers and pupils, as patent in syntheses of rare research on the effectiveness of practices in teaching to read and write in French (Feyfant, 2011; Talbot, 2012; Gausset, 2015; Écalte & Magnan, 2015).

A first national-scale study was launched in 2004 under the *Programme incitatif de recherche en éducation et formation* (PIREF—programme for the promotion of education and training research) in response to the consensus conference calling for a national project on the teaching and learning of reading (PIREF, 2003). It was conducted by a cross-disciplinary team that had attempted to characterize teaching practices in reading and mathematics classes and to determine their effectiveness at first grade (Sensevy *et al.*, 2007). The first objective was achieved, but not the second—the grain of analysis adopted to describe teaching practices (verbal interactions between teacher and student) could not serve to correlate practices with pupil progress.

Although didactic grounds for the effectiveness of teaching reading were not identified, this research stream did establish that the tasks proposed by the teacher, their type and their level of difficulty over the class-wide average play a decisive role in the quality of teaching delivery. “For research inquiry into teaching effectiveness, the task given to the students is a fundamental unit of analysis” (Sensevy, 2007, p. 351).

The Ministry of National Education’s General Directorate of School Education (DGESCO) moved in 2010 to commission a second study led in the Greater Lyon region and directed and supervised by Édouard Gentaz. This study mobilized a different methodological approach in an effort to patch up the gaps in previous research: the randomized controlled trial (Robert & Teillard, 2012). This approach, already dominant in the Anglosphere (Meuret, 2007) and slowly gaining ground in the

French-speaking world (Rey, 2014), compares an experimental intervention group, composed of teachers randomized to trial an innovative teaching device, and a control group, composed of teachers that continue to deliver their usual teaching. The difference in mean performances between the two groups of pupils is interpreted as the average causal effect of the intervention. The Lyon-based study showed that the first-grade pupils assigned to an intervention based on strategies borrowed from experimental cognitive psychology, including stepped-up instruction on grapheme-phoneme correspondences, did not learn to read any better than their peers (Dehaene, 2011; Gentaz *et al.*, 2013). This outcome finding was difficult to interpret, as the investigators had not designed any control for the practices implemented in-class, so we have no way of knowing whether the experimental instruction was significantly different to mainstream practice and how far the actual teaching practices of experimental-group teachers stuck to script.

It is in this context that we tabled a project to the French Ministry of National Education in 2012 to evaluate the effect of practices for teaching reading and writing on attainment of learning in first-grade pupils. The project reported here mobilizes the scientific paradigm of educational effectiveness research (Van Damme *et al.*, 2009) building on papers addressing “teacher effects” in elementary school (Attali & Bressoux, 2002; Nye *et al.*, 2004). In contrast to experimental approaches that force change in practices to better understand their impact, the goal here is not to transform the way teachers teach but to identify the didactic characteristics of the practices that prove most effective and most equitable. It is to fit this paradigm that we proposed to construct a huge cohort of teachers borrowing an array of didactic approaches.

We anticipate that our results will serve to inform strategic policy on the schooling system (*Centre d’analyse stratégique*, 2011) and on the content of pre-service teacher training and ongoing in-service teacher education. We also anticipate that our findings could help primary school teachers better frame and filter the choices open to them—or even make these choices intelligible to non-specialists, typically parents, which in the long haul should help defuse battling and bickering over methods.

The project was accepted by director-general Jean-Paul Delahaye in September 2012. It had been initiated by Roland Goigoux in autumn 2011. As lead research coordinator for the *Institut français de l’Éducation* (IFé—French Institute of Education) hosted at the *École Normale Supérieure de Lyon* (ENS Lyon—elite French school of research-training education), Roland Goigoux had sounded out colleagues via the *Association internationale de recherche en didactique du Français* (AiRDF—international association for didactic research in French). His first objective was to pull together a group of tenure-track lecturers representing an array of discipline specialisms, standpoints and schools of thought, including didactic-method theorists, linguists, psychologists, specialists in quantitative approaches in educational science, and more. Seventy-five university lecturers (including thirteen university professors and some of their PhD candidates) thus opted in, joined forces, and collaborated for four years. Their names are published in preamble to this long-form report, and they come from faculty teams hosted at universities across France—Amiens, Artois, Bordeaux, Caen, Cergy-Pontoise, Clermont 2, Grenoble Alpes, Lille 3, Lorraine, Montpellier 2 and 3, Nantes, Paris-Est Créteil and Paris 8, Rennes 2 and Bretagne Occidentale, Rouen, and Toulouse—Le Mirail. This research-group coalition set out the focal issue and methodology, conducted the field investigation, and processed their data which is under co-ownership terms.

His second objective was to pull together an even bigger group of field investigators to—simultaneously—observe the in-classroom practices of over one-hundred-and-thirty first-grade teachers. To be credible and designed with the scientific method needed to demonstrate causality in

social science, our study had to cover a very large sample of pupils and teachers at work in an array of school contexts. Direct observation of practice was a vital prerequisite to avoid the bias that declarative self-reported practice alone would fatally introduce. Over a hundred teacher trainers and educational advisors—some in work, others retired, but all purpose-trained for this investigational inquiry—thus joined the ranks of these academics to complete the observations over three full weeks. Their names also feature in preamble, with our warmest thanks.

Our study, led between September 2013 and June 2014 (first grade) and then in June 2015 (evaluations at the end of second grade), was co-financed by the French Ministry of Education (DGESCO) and IfÉ–ENS Lyon, followed by education research laboratory ACTÉ (Université Blaise Pascal—Clermont-Auvergne), with added support from the universities listed above. The bulk of our budget (around €200,000) went to payroll the students performing the individual or subgroup evaluations on the two-thousand-seven-hundred-and-five pupils in our sample cohort.

To study mainstream-curriculum classes in everyday situation, we reached out to teachers in fourteen different *académies* [regional education authorities], asking them to simply keep teaching as they normally would. The very first contact was couched in these terms: “Our objective is not to demonstrate the superiority of any one method, but simply to identify—with no preconceptions whatsoever—the characteristics of the practices that prove most effective and most equitable. For this purpose, we are looking for experienced primary school teachers who are teaching to mainstream curriculum but interested in finding out more about what actually shapes the quality of their pupils’ learning. We would therefore be delighted if you would consent to participate in our study by letting one of our field-researchers—who would be the same person all year through—come in and talk with you, analyze your pupils’ progressions, and observe your teaching practice.”.

One-hundred-and-thirty-five teachers answered our call for participants and opted in—some seeking confirmation of the didactic and instructional theory that they staunchly defended, while others—the majority—simply wanted to know whether the actions they adopted actually work. One-hundred-and-thirty-one were observed by the same investigator or pair of investigators (one-hundred-and-ninety-two field-investigators in total) from the start to the end of the year. Our thanks go out to all of them, even if we have not published their names here to maintain the anonymity and confidentiality central to our ethical social research practice policy¹.

Finally, we also thank the regional education authority directors, national education authority inspectors, and local education authority constituencies who facilitated this investigational inquiry.

¹ All information collected and published in this report has been de-personalized and de-identified. Each researcher has access to the full dataset, made available via a secure website, as per the principle of data co-ownership and co-sharing that our alliance signed up to in February 2012. The lead project manager, engineer and statistician are the only people in possession of the codes needed for re-identification of classes. However, each field-investigator has been given the code number of their class they followed, and equally, each teacher can access all information identifiable to them.

A.2 Focal issue

A.2.1 Research questions

The purpose of this research was to contribute firm scientific evidence on the effects of different didactic-strategy interventions in teaching beginning reading and writing. The aim was to identify which contents, tasks and instructional strategies work best for pupils' reading and writing education at different periods of the school year, and to find out which trajectories of pace, depth and mix were differentiated for different pupil audiences targeted. We also aimed to resolve the issue of whether the content defined in statutory curriculum programmes should be taught sequentially or simultaneously, isolatedly or integratedly, explicitly or by straight reiteration. Another question we wanted to address was whether practices that prove effective share common traits, and if so, which. Finally, we also aimed to resolve the issue of whether the factors of effectiveness identified in experimental research or earlier action research translate up to natural context, i.e. in contexts not transformed by the research frame itself.

In order to single out the indicators capturing didactic strategies mobilized by the teachers and thereby construct the variables serving to underpin our subsequent analysis (see A.3.1.4), we discerned five subsets of questions covering the teaching of:

- the alphabetic system and word identification processes,
- comprehension of written print,
- writing,
- studying language,
- cultural literacy.

This breakdown—a mainstay of French language didactics—articulates issues for scientific research (Nonnon & Goigoux, 2007; Daunay, Reuter & Schneuwly, 2011; David & Morin, 2013; Kervyn & Brissaud, 2015) with the concerns of people on the classroom frontline (teachers, instructors, leadership and support), because it also fits with the areas of competency defined in curriculum programmes and with the goals and objectives assigned to teachers.

Didactic variables may be the main target variables of this study, but they are by no means the only ones, as other more 'instruction-centric variables', which are more cross-cutting and less content-area-focused, may also be factored in based on earlier research in educational science as good candidates for explaining what produces differences in effectiveness (Dumay & Dupriez, 2009).

We have discerned three subsets of variables covering:

- the explicitness of the teaching delivered,
- differentiated instruction and support to struggling pupils,
- classroom climate and pupil engagement in tasks.

These eight subsets of questions on teaching practices form the focal issue tackled, which we ground and detail in the sections that follow here.

They are presented in full, along with the methodology adopted (see A.3) and the three-part write-up of the study results, in the full report posted up online on the IFé website on 15 March 2016. <http://ife.ens-lyon.fr/ife/recherche/lire-ecrire/rapport/rapport-lire-et-ecrire>

- Part one (section B of the full report—descriptive analysis: what do the pupils know?) reports on the knowledge and skills of the pupils evaluated at three timepoints (start of first grade, end of first grade, and end of second grade [or elementary-school—year one in French-system terms]) in order to gauge their progression.
- Part two (sections C and D of the full report—descriptive analysis: what do the teachers do?) reports on the practices of first-grade teachers.
- Part three (section E of the full report—inferential analysis: finding causality) reports on the influence of teachers' practices on pupils' learning.

Here we give only an executive brief reporting our key findings concerning the influence of teachers' practices on pupils' learning (section B, page 38)

A.2.2 Questions on the teaching of the alphabetic system

At the time we defined our research questions, primary school education in France was governed by the national curriculums issued in 2008 which had put a top-down end to 2005–2006 polemics over the methods for teaching reading. The guidance documents published in the Ministry of National Education's official bulletin (*BOEN*, special issue #3 of 19 June 2008) stipulated: "the national curriculums for primary school education define, for each area of learning, the knowledge, skills and competencies to be attained in staged cycles education; they set out annual key-stage goals to help frame and structure a progressive curve of learning in French and mathematics. They do however leave primary school teachers free to choose their teaching methods and instructional approaches, in a demonstration that the system trusts its teachers to adapt the teaching delivered to their pupils".

The then minister of education Xavier Darcos, in his policy memo to the teachership, had stressed the necessity of having free-choice method: "This instructional freedom serves your school, your purpose, your missions: it gives you latitude to adapt the progressions targeted to your pupils' needs; it helps you to inclusively blend equal access to instruction with the inherent in-classroom diversity of children and contexts; it cements our recognition of the qualifications, expertise and professionalism our teachers and head teachers are demonstrating day-in day-out" (p. 1).

The preface to these new curriculums did carry one piece of methods-related guidance that cramped the principle of instructional freedom, as it pruned "structured and explicit" teaching, although without any concrete pointers on what these adjectives were meant to mean. The cruelly short paragraph on teaching reading and writing in first grade gave no indication as to how pupils were supposed to gain a commanding grasp of the alphabetic system, stating simply that "pupils practice learning to decode" and that "this practical learning progressively helps the pupil read faster and more fluently"³ (p. 2).

Four years down the road, in January 2012, i.e. eighteen months after the start of our field investigation, the key-stage goals proposed to teachers as guidance on instructional progressions⁴

² <http://www.education.gouv.fr/bo/2008/hs3/lettre.htm>

³ http://www.education.gouv.fr/bo/2008/hs3/programme_CP_CE1.htm

⁴ http://cache.media.eduscol.education.fr/file/Progressions_pedagogiques/78/6/Progression-pedagogique_Cycle2_Francais_203786.pdf

stated that by the end of first grade, in the area of command of alphabetic system, pupils were expected to:

- name the letters of the alphabet, in order;
- distinguish the letter and the speech sound it represents; know the correspondences between letters and sounds in simple graphemes (e.g. *f* ; *o*) and more complex digraphs and trigraph (e.g. *ph*; *au*, *eau*);
- know that the syllable is composed of one or more graphemes, and that a word is a blend of one or more syllables; be capable of segmenting a word into these lexical units (graphemes, syllables);
- easily read the words studied;
- decode unknown regular words;
- easily read the most familiar words they encounter (dubbed sight words);
- read aloud a short text composed of familiar words, clearly and fluently and using the right punctuation;
- choose and write their own simple words, using the right letter–sound correspondences. (MEN, 2012, p. 2).

Put briefly, at the time of our inquiry, first-grade teachers were free to forge their own approach-method for teaching reading and writing, on the caveat that they aimed to attain the above-listed competencies.

As tenure-track lecturers and teaching fellows, many of us had been engaged in their initial pre-service and ongoing in-service teacher education, so we were well aware of their doubts and concerns. We knew that findings from scientific research, while offering solid pointers, still fell short of addressing their central concerns, particularly questions over when and how to schedule the study of grapheme-phoneme correspondences. Questions like should explicit phonics instruction be started in the early days of first grade? Can the teacher rely on the rote-learned memorization of whole words, introduced before the children know how to decode them? And moving forward, what space should be left for analytic phonics and synthetic phonics? Which linguistic units should be given first focus? In what order should they be studied? At what pace? Should we spend the bulk of teaching time learning to decode at the start of the year? How far can encoding activities facilitate decoding activities? Is it better to start by studying the phonemics before transitioning to their transcription as graphemics (encoding) or, conversely, to start by graphemes first and then their conventional values as sounds (decoding)? How do you choose or draft text resources to support teaching to decode? Is it a valuable exercise to show pupils texts they are not yet equipped to decode? Etc.

These professional-practice questions have become our research questions because they directly resonate the scientific debate we are engaged in here. The nucleus of the problem is that even though an international consensus began to emerge in the early 2000s around the idea that early, explicit, systematic teaching of decoding facilitates initial learning of reading–writing⁵

⁵ This consensus had already influenced the French curriculum programmes drawn up in 2002, which told teachers to help their pupils memorize the grapheme–phoneme mappings, enable them to firmly structure this information, regularly go back over and consolidate the mappings, and systematically teach for “memorizing the main syllabic onset-rime groupings of vowels and consonants in the various combinations possible” (MEN, 2002, pp. 44–45). The 2002 programmes of study specified that this teaching was to further revolve around “analyzing whole words through their smaller units based on knowledge already learned” (idem), using, for example, analogy phonics (using the “*pa*–” syllable in “*papa*” to segment and

(Sprenger-Charolles & Colé, 2006, in the wake of the National Reading Panel Report, 2000), the format for this instruction is still a controversial issue, as it has been under-studied in French language. Illustrating this gap, no study had compared the effectiveness of didactic approaches qualified as grapheme-first (based on studying the sound that letters make) *versus* phoneme-first (based on studying how sounds can be written down), or put another way, approaches that put decoding first (letter-to-sound correspondences) *versus* approaches that put encoding first (sound-to-letter correspondences). The above-cited English-language research, which corroborates the conclusions of Goigoux (1993, 2000) and Braibant & Gérard (1996) in French language, only found evidence for superiority of systematic phonics instruction (whatever the format) over unsystematic or no-phonics instruction. The National Reading Panel (2000) final report explicitly concluded that there were two approaches possible: “synthetic phonics programs—teaching pupils explicitly to convert letters into sounds (phonemes) and then blend the sounds to form recognizable words” and “analytic phonics programs—teaching pupils to analyze letter–sound relations in previously learned words to avoid pronouncing sounds in isolation” (NRP, 2000, p. 93).

In short, the question of programme-of-study schedules for alphabetic system instruction was severely under-documented, especially in French language instruction, and has remained so since (Nonnon & Goigoux, 2007; Feyfant & Gausse, 2007; Écalte & Magnan, 2010). As touched on in the introduction, only researchers allied with Édouard Gentaz in 2010–2011 had attempted to tackle the issue by comparing progress trajectories in two groups of pupils: an experimental intervention group proposed a grapheme-first approach based on a programme-of-study schedule for grapheme-phoneme correspondences instruction established by Liliane Sprenger-Charolles (2011) according to linguistic criteria, and a control group delivered the usual instructional practice leaning towards phoneme-first without excluding the orthographic memorization of whole words. There was no difference in effectiveness between these two groups, and no firm grounds to suggest either approach was superior to the other (Dehaene, 2011; Gentaz *et al.*, 2013).

This prompted us to revisit the issue here, in our own design, and examine the different formats for scheduling alphabetic system instruction as observed in French first-grade classes in an attempt to find out whether certain models work better than others. We wanted to know whether explicit and early-on instruction of grapheme-phoneme correspondences was as actually effective as we surmised. We wanted to know whether results established in English-language instruction translated and transposed into French-language instruction, and which of the instructional practices would fit these criteria (early and explicit instruction) which are operationalized in highly variable ways between different countries with different teachers’ backgrounds, pathways in education and instructional cultures. We too thought it axiomatic that pupils should be given early experience of learning with the alphabetic principle discovered in preschool, but we premised that it could be delivered from within an array of instructional approaches, including approaches designed around strong early teaching for learning to encode. The thinking behind this premise has roots in Célestin Freinet’s legacy and articulates with innovations orchestrated by the *Institut national de recherche pédagogique* (INRP—French agency for education and instruction research). We also thought that this instruction needed to focus on a significant number of some of the most common correspondences in the French orthography to enable the pupils to swiftly succeed in decoding the written texts proposed by their schoolteacher. We took the stance that a swiftly *tempoed* schedule

decode the new word *Paris*). Put simply, analytic phonics and synthetic phonics were deemed complementary as activities for teaching to read and write, and the schoolteachers were encouraged to tackled them together.

of instruction in grapheme-phoneme correspondences would likely benefit learning, as it could lend greater cognitive clarity to pupils (Downing & Fijalkow, 1984) and enhance their self-learning ability, while at the same time cutting short room for discouragement and groping in the dark (Goigoux, 1993).

This is a stance that has been defended by our community of practice for nigh-on two decades (Gombert *et al.*, 2000) and reinforced by the conclusions of the 2003 consensus conference (PIREF, 2003), but never actually put to test in a French-speaking country. We therefore took this opportunity to test its validity, for example by recording the number and kind of grapheme-phoneme correspondences explicitly taught in the first period of the school year and calculating the decodable fraction of the code-in-print resource material⁶. We wanted to find out in which context-set this stance holds as grounded and whether all pupils would be set to benefit regardless of their initial level. We also wanted to know whether *tempo* of curriculum delivery influenced quality of learning directly or else in interaction with other didactics variables, such as writing, cultural literacy, etc. (Bucheton & Soulé, 2009; McGill-Franzen, 2010).

If no comparative study on reading “methods” has managed to establish the superiority of any one method, it is not because all instructional practices pan out as equally effective, but because the “method” variable, which is too coarse and loosely definable, is simply not relevant as variable for identifying the drivers behind differentiated practice-related effects (Fayol & Goigoux, 1999). It is also because there is a world of difference between assigned work and actual work (Tardif & Lessard, 2000; Lantheaume, 2007). That said, the methodology principles underpinning the various approaches will undeniably translate in concrete terms through contrasted words-for-learning proposals in the textbook readers. We thus set out to address the influence of using—or non-using, for those teachers that elect to go without—the assigned primary-school readers on teachers’ practices and, therefore, effectiveness. For this purpose, we had to examine the actual teaching practices of schoolteachers using the same reader (were they much the same or not?) and map out the programme-of-study schedules for written-code instruction associated with use of these readers—readers that we have collapsed into three categories based on whether the author’s approach is embedded, synthetic or analytical phonics.

In short, our main questions were these:

- 1 do the number and kind of grapheme-phoneme correspondences taught at year-start have an effect on pupils’ code performance outcomes?
- 2 do pupils make faster progress when the route in is graphemic-first or phonemic-first?
- 3 do pupils make faster progress when they spend a large amount of time on encoding tasks? If so, is progress driven by encoding familiar graphemes/phonemes chosen by the teacher or by trial-and-error through writings chosen by the pupils (“invented” or “trial-and-error” spelling)?
- 4 does increasing the decodable fraction of code-learning resource material have positive effects on pupil performance outcomes in code tasks?
- 5 does choice of reader influence programme-of-study schedule for written-code instruction and, hence, effectiveness of the teaching?

⁶ The hypothesis tested was that the teachers employing a large fraction of readily-decodable print (sentences, text) in the route to decode words would be more effective than teachers working with less-decodable print.

A.2.3 Questions on the teaching of comprehension

The hypotheses on comprehension are built on two observations. The first is that even though children have typically developed a high level of communicative competence when they start first grade, they do not yet have all the language skills in place that they need for written text comprehension. The second is that it is not as simple as just teaching them to decode the written word so that their language comprehension skills will kick in and enable them to understand written texts just as well as they understand oral language. This, according to Oakhill, Cain & Elbro (2014), *“is a misconception because it ignores the fact that written texts are, in important ways, different from spoken interactions [...] and written texts typically require memory abilities and other cognitive skills that are not so crucial in understanding everyday interactions”* (p.5). Children’s oral language comprehension skills are built on processing discourse-context information and nonverbal communication clues whereas writing skills are characterized by the absence of referent and interlocutor. Writing thus requires higher-register lexis, with a very different syntactic structure and a new kind of textual cohesion and discourse organization, so children need time and exposure to progressively internalize these new patterns (Bronckart, 1985). The process of writing, which unlike spoken language leaves a permanent physical trace, also implies an ability to self-regulate reading (modulating pace and pauses, backtracking, and detecting incoherency, etc.) that is on a whole different level to the ability needed to self-regulate oral dialogue.

For first-grade children, learning to read is not just about learning to decode, it is also about becoming capable of making sense of text they read by mobilizing substantially the same competences as those required to comprehend text they hear, such as when parents or educators read to them (Hogan, Adlof & Alonzon, 2014). Nevertheless, there is a large evidence base—often downplayed by proponents of a simple code-oriented view of reading (ONL, 1998)—showing that listening comprehension skills in young children at the end of nursery school are excellent predictors of effective writing comprehension achievement across the developmental span, even years later (Hirsh, 2003; Hoff, 2013; Durham *et al.*, 2007; Kendeou *et al.*, 2005; Oakhill, Cain & Bryant, 2003; Paris & Paris, 2003; van den Broek *et al.*, 2005).

These competences can be put into four categories:

- language processing-specific competences: syntax, lexis, morphology, processing anaphora...
- specific text processing (reading or listening) competences: features and structures of texts, as narrative, explanatory, prescriptive (etc.), culture and world knowledge...
- cognitive competences: integration of successive informational items, construction of a mental model, reasoning, memorization...
- strategic competences: inferencing, control of comprehension...

Several studies show that these competences are acquired early, as most 4-year-olds are able to construct a mental model of text they hear, maintain a coherent version of the story, make inferences to piece information together, fill in the gaps left open in the story, and answer inferential questions (Kendeou *et al.*, 2008, 2009; Tomkins, Guo & Justice, 2013; Florit, Roch & Levorato, 2014)—all competences that, if in deficit, will block a child’s comprehension even when they can identify the words without problem. This is why a number of researchers argue that if we want to give all children a chance to acquire these skills, then we have to teach them explicitly, and early on, without waiting for them to know how to decode. The experiments led to date have yielded very

encouraging findings in both children on a typical development trajectory and children with language development deficits (Catts *et al.*, 2015; Clarke *et al.*, 2010; Sénéchal & Lefevre, 2001; Snow, Burns & Griffin, 1998). This core teaching is doubly vital as preschool-age text comprehension predicts later-stage reading comprehension (Griffin, Hemphill, Camp & Wolf, 2004; Kendeou *et al.*, 2008, 2009; Paris & Paris, 2003; Storch & Whitehurst, 2002), and doubly warranted by convincing evidence that the two subsets of competence (code and comprehension) actually develop relatively independently (Cain, Oakhill & Elbro, 2014; Cutting & Scarborough, 2006; Kendeou *et al.*, 2005, 2007, 2009; Muter, Hulme, Snowling & Stevenson, 2004; Savage, 2006). This all means that, on the didactics front, there is absolutely no good reason for first-grade teaching to break the trajectory of listening comprehension started and engaged in nursery school. In other words, first-grade teachers “must not only be attentive to the cognitive processes underpinning skilled decoding, but equally attentive to the processes needed for comprehension” (Oakhill *et al.*, 2014, p. 9), because if they are not, then the children arriving underequipped at first grade will very likely end up unable to comprehend what they are perfectly able to decode.

A.2.3.1 *Early and persistent differences*

The very nature of the processes inventoried above explains why children from low-SES homes struggle hardest with comprehension (Hoff, 2013; Daussin, Keskpaik & Rocher, 2011). There are several environmental factors that converge.

The first stems from language socialization experiences on the family side. Studies by Hart & Rissley (1995) in the USA and Lenormand *et al.* (2008) in France, for example, show that in many families from low-SES homes, the language employed counts fewer complex sentences and less lexical variety than in families from high-SES homes, and Hoff (2013) and Huttenlocher *et al.* (2007) also observed that conversations in low-SES homes tend to be shorter, hampering the development of lexical and grammatical competences in both speech input and speech output. Sim *et al.* (2014) also showed that “shared reading” was used both differently and less often in low-SES homes. In a study comparing family-side language socialization practices between contrasted social backgrounds, Heath (1982) highlighted that these differences emerge as much in the kind of picturebooks read as in the ways they are shared with the preschooler children. In France, the study led by Bonnéry & Joigneaux (2015) arrived at the same conclusion: the way parents share reading at home is part of what “embodies children from very early age with a dispositional inheritance that they will take to school as their educational capital” (p.24). A large number of studies (for a review, see van Kleeck, 2008) shows that in higher-SES families, although parents start conversation around reading on the very literal level with simple and concrete elements, they will steadily, over the nursery-school years, encourage their children to make more inferences grounded in mobilizing higher-level reasoning. Early exposure to stories, reading and conversation about books promotes the development of reading skills. Not all children begin primary school on an equal footing because they do not all experience the same preschool practices, and, as we saw earlier, some practices are better than others.

The second factor stems from teaching practices employed at nursery school. The National Education Inspectorate report (2011) sheds light on this issue. It finds a lot of reading getting “offered” in nursery school but very little effort to work on comprehension. The authors are critical of this gap: “Class schedules regularly hold their promise: children in nursery school are read a lot of stories. But the vast majority of these same schedules make no mention of organizing interventions on their comprehension of the texts heard—neither the language of the storylines nor the language

arts connecting up to the pictures that they are almost always illustrated with” (p. 131). This makes it no surprise to find that, on this point and many others, nursery school fails to bridge social inequality gaps. Not only do inter-individual differences in language skills emerge early, but pupils who are bottom of the class on comprehension at four years old tend to stay there—in France as elsewhere (Torsch et Whitehurst 2000, 2001; Catts *et al.*, 2006; Durham *et al.*, 2007; Elwèr *et al.*, 2015; Nation *et al.*, 2010).

That said, the scientific literature is rife with resources to assert that this should not be a losing battle. The next subsection discusses why.

A.2.3.2 *Some practices more effective than others*

Effectiveness studies in comprehension skills into research can be synthesized into a list of instructional practices that have demonstrated proven effectiveness for improving comprehension skills. It comes as no surprise to see that effective schoolteachers explicitly teach comprehension. For each text, effective schoolteachers start out very proactively guiding the construction of meaning, then progressively take more of a back seat to hand more responsibility to pupils on this activity (Duke *et al.*, 2011). They provide multiple encounters with a large volume and range of text (Taylor, Pearson, Clark & Walpole, 2000). They provide motivating texts and contexts for reading (Turner & Paris, 1995; Williams *et al.*, 2009). They focus pupils’ attention on text structures: they teach pupils to construct a coherent local mental model, by making logical or chronological connections between the sentences and pieces of information in the text, and a coherent macro (whole-text) mental model, by encouraging them to integrate the sentences and pieces of information into a coherent whole. They also get pupils to learn to connect said information to their own knowledge. They make learning and memorizing lexis a constant ongoing goal (for a background review, see Goigoux & Cèbe, 2013). They proactively build world knowledge (Hirsh, 2003). They organize plenty of discussion to talk about the text (Murphy, Wilkinson, Soter, Hennessey & Alexander, 2009). They ask high-level questions during open discussions and foster high on-task engagement and pupil participation rates (Murphy *et al.*, 2009). They integrate—not isolate—reading and writing (Morrow, Tracey, Woo & Pressley, 1999). They deliver differentiated teaching (Connor *et al.*, 2009), observe their pupils and regularly evaluate their rates of growth in comprehension (Duke *et al.*, 2011).

This list warrants being completed by studies focused more specifically on primary-school populations. The review by McNamara & Kendéou (2011) prompts us to add that effective schoolteachers give pupils all sorts of opportunities to practice paraphrasing and reformulation, regularly ask them to make inferences (bridging inferences, causal inferences), encourage them to make educated guesses, encourage them to connect in-text information to other in-text information and to their own cognitive psychology and world knowledge.

All this converges to explain why there is a whole scholarship defending the idea that if we want to teach primary-school pupils how to comprehend text and mitigate social background-related differences in efficient comprehension, then we have to include specific teaching for all of these comprehension skills right from nursery school and continue it, uninterrupted, into primary school/first grade. The study led by Snow, Barnes, Chandler, Goodman & Hemphill (1991) substantiates this policy, as it found that at the start of elementary school, pupils from low-income backgrounds make significant progress if they get real comprehension instruction over the following two years, whereas only 25% will make significant progress if they get only one year of real

comprehension instruction, and they will make no progress without real comprehension instruction. The study also found that schooling influence was particularly acute for complex cognitive tasks.

A.2.3.3 *Research questions*

This literature review leaves us ready to refine our research questions in the field of comprehension instruction. To identify which in-classroom practices are most effective, we first have to evaluate the pupils' initial levels of competence in lexis, syntax and listening comprehension (see A.3.3.1) and then come back and evaluate their final levels of competence in independent reading comprehension and in listening comprehension (see A.3.3.2), which will give us a frame in which to evaluate the influence of the teachers' practices on the pupils' progress in both areas of competency.

To describe the features of instructional practices and insert them into our explanatory model (see A.3.1), we will investigate whether the main hallmarks of quality teaching identified in our earlier synthesis of the research can be found, in some measure, in the practices of first-grade teachers in our sample. We will thus endeavour to inventory all the teaching tasks allocated to comprehension over the three school terms of the year, with the times on tasks and the kinds of tasks. We will observe, for instance, how far pupils are encouraged to: explain or reformulate the meaning of a sentence or text; recount a mental model; produce a narrative recall or explanatory text recall; make implicit information explicit; propose, debate or negotiate an interpretation, etc. (see A.3.4.1.1).

In short, our line of inquiry is to find out:

- whether time allocated to teaching reading and listening comprehension has an impact on pupils' progressions;
- whether regular and intensive comprehension teaching throughout the year has positive effects on reading and/or listening comprehension;
- whether—instead—focusing the year programme first on code instruction before moving on to comprehension later is a strategy that works;
- whether teaching comprehension isolated from code works better than teaching that mixes the two;
- whether intervention based on an integrative textbook leads to more comprehension-oriented teaching than with a code-oriented (synthetic or analytical phonics) textbook;
- whether teachers that arrange a lot of opportunities for conversation around reading are more effective than other teachers;
- whether pupils that have learned to explain, reformulate, paraphrase or summarize make more progress in comprehension;
- whether time allocated to oral dialogue on written texts has the same influence as time spent on comprehension-oriented writing tasks on written texts;
- whether teachers that teach their pupils to learn and memorize lexis are more effective than other teachers;
- whether teachers that combine regular and intensive comprehension teaching throughout the year with study of language (including lexis) are more effective than other teachers;
- whether the number of texts listened to and read has a beneficial influence on outcomes;
- whether pupils that get to hear more stories read by their teacher or decoded by themselves make more progress than other pupils.

These questions will be studied by observing the differentiated effects of instruction on learning trajectories in pupils starting out with different initial levels of competence. We will thus compare the progressions of initially poor comprehenders against initially good comprehenders, but also against pupils from low-SES backgrounds and pupils schooled in France's 'priority education network' schools for struggling communities.

A.2.4 Questions on the teaching of writing

How and how far do teachers' writing practices influence pupils' reading learning (code and comprehension) and writing learning itself? This wider focal issue leads us to address several questions—what are the writing tasks prescribed to pupils in the classroom (handwriting, copying, dictation, composition)? Which of these tasks have the most influence on reading and writing learning? Is it the time-on-task factor that has an effect on learning? If there is an evidence-based effect of time on certain tasks, does time on these tasks have floor effects and ceiling effects on learning? To inform complementary analysis, are there types of linguistic unit worked on in writing tasks (letter, syllable, word, sentence, text) that have more influence on learning? Finally, which combinations and types of sequences or linkages between writing tasks have an effect on learning?

The major hypotheses on these issues are that reading and writing are tightly intertwined and that early-starting, regularity, volume and variety of writing practices all positively influence learning to read and write.

Studies on initial teaching to read and write have started to emerge how writing practice contributes. Several papers have clearly highlighted that writing informs the development of rethinking print, the development of reading mechanisms (Reuter, 1998; Fitzgerald & Shanahan, 2000; Écalle & Magnan, 2002) and literacy development as a whole (reading–writing) (Read & Treiman, 2013). Most of the texts compiled by Rieben, Fayol & Perfetti (1997) support the idea that the processes mobilized in learning to read and the processes mobilized in learning to spell are tightly related (Ehri, 1997, 2000)—an idea further consolidated in the *Handbook of children's literacy* (Nunes & Bryant, 2004) which extends this strong relationship to encompass text production.

Beyond this suite of questions, based on early quantitative processing of the dataset presented further down, we take our investigation further on to the characterization and conditions of effective didactic interventions by teachers during writing tasks. Our analytical strategy is essentially based on processing the data captured in video recordings to serve an in-depth study of the teacher's actions during writing tasks: choice of support material for writing and the way the teacher gets pupils on-task, regulates and scaffolds their activity or differentiates his/her forms of intervention... the goal being a model of the teacher's educational competence in pupils' writing production tasks.

In short, our line of inquiry is to find out whether (and under what conditions) early-starting, regularity, volume and variety of writing practices positively influence learning to read and write.

A.2.5 Questions on the study of language

Right out of early childhood education and into the basic skills stage of elementary school, the official 2008-year national curriculum programmes set out the priority objectives for first-grade schooling as "reading, writing and the French language [...] with in-class acquisition in these domains given permanent ongoing focus in any and every class activity being led". The texts specify that reading and writing education "is to include the first steps on the path towards grammar and

spelling”, which looks to endorse exploring the grammar system as an integral part of early literacy education. These prescriptions mark a shift from the policy from the prescriptions issued when key-staged education cycles were ushered into primary school education (1995, 2002) setting more limited objectives on metalinguistics education at the end of second grade: gender and number agreement in noun phrases and subject–verb agreement in regular sentences. In 2008, the progressions for elementary school, designed based on curriculum programmes and schooling for the core foundation standards, list the skills that pupils are expected to have acquired⁷. These progressions are to enable teachers to frame and structure a progressive curve of learning for each year of the education-stage cycle (BOEN, 19 June 2008) yet they do not stipulate whether observation of how language works is supposed to be introduced, – and if so, how, – with any specific metalanguage. These earlier-track requirements raise a number of questions: are we to teach the study of language in first-grade classes, and if so, how? What linkages are established between study of language and teaching of literacy in the practices observed? What place is given to introducing early exploration of grammar, in what form is it introduced, and what effects does it have on pupils’ learning outcomes?

Tackling these questions led us to posit six hypotheses, constructed into two strands—one strand on the effect of time spent studying language on pupil’s learning to read-and-write, the other strand on the effect of the metaterms⁸ used in-class.

In short, our line of inquiry is to know whether, as we expect to find:

- time spent studying language has a positive effect on pupils’ results in learning to read-and-write;
- time spent studying language as a main-aim and/or secondary-aim task has a positive effect on pupils’ results in learning to read-and-write;
- time spent studying language in tasks coupled to comprehension or writing tasks has a positive effect on pupils’ results in learning to read-and-write;
- time spent studying language throughout the year or just at the end of the year (time-on-task “spaced out from early on” versus “packed together from late on”) has a positive effect on pupils’ results in learning to read-and-write;
- using written metalanguage terms (official or ‘hand-made’) has a positive effect on pupils’ results in learning to read-and-write;
- the presence of fundamental metalanguage terms—sentence, noun, verb—will show up in the most effective classes.

At this stage of our inquiry, the results and findings reported here essentially cover the question of the effect of time spent studying language on children’s learning to read-and-write. Quantitative

⁷ For first grade, pupils are expected, for instance, to be able to:

- Recognize nouns and verbs and distinguish them from other words; distinguish the noun from the article that goes before it; identify the article; beginning with pronouns: know how to use the subject pronouns orally.
- Discern and argue markers of gender and number: the ‘s’ marking the plural form of noun, the ‘e’ marking the feminine form of the adjective, the -nt endings of first-group verbs in the present indicative.
- Start independently using markers of gender and number (plural form of the noun, feminine form of the adjective, the -nt ending of first-group verbs).

⁸ The noun *metaterm* and its equivalent *metalinguistics* term refers not just to grammatical categories such as verb, singular, plural, etc., but also to mainstream-curriculum lexical categories forged by schoolteachers, such as “little word”.

processing analysis of the metaterms compiled from in-class sessions has pointed to some early insights, but study of the effect of metalinguistic discourse is still in the pipeline.

A.2.6 Questions on literacy enculturation

What literacy enculturation practices (or “enculturational practices”) do teachers propose and what are the impacts on children’s development of reading-writing skills? There is a whole stream of French-language research (Bonnéry, 2012; Devanne, 2006; Fijalkow, Pasa & Ragano, 2006; Frier, 2006; Giasson 2007; Tauveron, 2002; Terrail, 2013) highlighting the importance enculturational practices have in helping children build the habit of reading in first grade, but none has ever stopped to scientifically look at how the practices actually link up to pupils’ progress. The objective here, then, is to inquire into this process of enculturation in school and its contribution to the success of in-school learning.

This section sets the objectives, theoretical scaffold, and hypotheses guiding our line of inquiry.

A.2.6.1 Talking points on the concept of literacy enculturation

Becoming reader–writer involves much more than learning a set of technical skills in school. It also requires an enculturation process, which can be defined as “appropriation and familiarization with the culture of writing—its printed works, its linguistic codes, its social practices” (Goigoux, 2003, p. 2). This enculturation process lends meaning to linguistic instruction and enables children to “discover the powers of agency and thought that command of written language brings (schoolteachers talk about the construction of reader “status” or “posture”)” (*ibid* p. 2).

There are two ways to grasp enculturation: as a “process by which a person or group assimilates the culture of another group” and as “adaptation of a person to a surrounding culture” (Chauveau, 2011, p.156).

Literacy enculturation thus works on two fronts. Chauveau adds that, given how “there is simultaneously an exteriorization process [...] and an interiorization process [...], learning to read involves at the same time reaching out into world of the written word, discovering it, inhabiting it, practicing it (the pragmatics dimension), and breathing it in, making it his own, transforming it into new attitudes, new ways of thinking, and new self-reflective dispositions (the psychic dimension)” (2011, p.156).

This cultural integration will therefore kick in very early on, scaffolded by the child’s interactions with the readers–writers community, through an array of activities, settings and situations.

Focusing in on the literacy enculturation process, there is an observable shift between the community frame (social referencing practices, rituals, etc.) and the individual frame (appropriation). In other words, children do not enculturate literacy without outside influence— which is precisely what makes mediation – especially in school – and the enculturational practices operationalized by the adult mediator or ‘facilitator’ (teacher, librarian, author, volunteer...) so crucially important.

A.2.6.2 Dimensions of literacy enculturation

The enculturational interface with the world of writing carries several dimensions that, although articulated together, can be studied separately:

- a literary dimension: constructing a first literary culture (Devanne, 2006; Tauveron, 2002; Terwagne, 2012);
- a cultural and anthropological dimension: platforming narrative immersion. Here, story is seen as the most powerful matrix of discourse, as it is through story that all other forms of discourse emerge, in an anthropological view of culture (Bruner, 2002; Le Manchec, 2005);
- a social and community dimension: joining the community of readers, understanding the codes and social uses of writing (André, 2003; Chauveau, 2011; Privat, 1993);
- a psycho-affective dimension: constructing the habit of reading widely and often, constructing a personal relationship with the world of writing, making room for subjectivity (Petit, 2002; Terwagne, 2012). Enculturation through writing is operationalized through a process of appropriation by the subject as reader (Langlade & Rouxel, 2004), which goes hand in hand with the development of reader postures (Bucheton, 2006). Enculturation cannot happen without appropriation—so not without development of the child’s own personal relationship with the written word;
- a cognitive dimension: learning to think in and through writing, which mobilizes a twin approach of distancing and reflective representation (Goody, 1979; Lahire, 1993; Bucheton & Chabanne, 2002; Terrail, 2013).

These component dimensions of literacy enculturation also emerge in scholarship on the “relationship to writing” (Barré-De Miniac, 1995, 2000) where *relationship* “designates the multiple affective, cognitive, social and cultural ties connecting [the student] to this activity” (Colin, 2014). How these ties are constructed determines how the pupil will lend meaning to what is learned at school, and so the role of the school in the construction of this relationship to writing becomes doubly pivotal to counter the weight of social determinism on certain pupils’ shoulders (Rochex & Crinon, 2011).

A.2.6.3 Hypothesis

The route into writing sets out as an enculturation process (Goigoux, 2003), where the child first has to discover the different literary, cultural, social, cognitive, and psycho-affective dimensions of writing, which we can conveniently crystallize in terms of relationship to the ‘written word’ (Barré de Miniac, 1995). The process plays a significant role in children’s development of reading-writing skills (Pasa & Ragano, 2008).

This enculturation process requires mediation from competent peers, or “literates” (Chauveau, 2011). It develops first in the home, in a way that is informal and unstructured, which may explain why different family-side cultures translate into different degrees of familiarization with written language in pupils starting out in school. (Lahire, 2008).

The process of gaining cultural literacy then continues way into first grade through teaching practices that strongly articulate the cultural and technical direction of learning (Fijalkow, 2000) via a frame of mediation associated with “effective” interventions, which has already been charted in various research work (Frier, 2006; Le Manchec, 2005).

This is our rationale for investigating whether classes that feature instruction specifically on cultural literacy help foster a positive relationship with the ‘written word’ and contribute to across-the-board success in every domain of pupil’s learning to read-and-write.

We expect to find a positive impact of enculturational practices in every test across the board at the end of first grade (code, comprehension, writing), particularly for those pupils that were less familiarized with written language when initially joining first grade.

To test this hypothesis, we propose to scan the practices observed for those we see as enculturational and, for there, to categorize and characterize the different classes in our sample on the basis of these practices. To operationalize this enculturation construct, we borrowed the Chauveau's model (2007) of four strands (or 'observables') for capturing the culture of writing in the classroom:

- the text-type: functional text (to name or list), informational text, fictional text, literary text, scientific text, prescriptive text (what to do, how to do it), etc.
- the object-media: different types of books, newspapers, textbooks, magazines, albums, encyclopaedias, logbooks, posters, letters, brochures, etc.
- the place-space: the reading corner, the library, museum, reading or writing workshop, etc.
- the peer-literates: the reading and writing practitioners, the 'users' of the writing.

Working to this initial categorization scheme will enable us to introduce variables into a multilevel model analysis so as to measure the—assumedly positive—effects of enculturational teaching on writing development skills. The categorization scheme can be found in section C.6.

A.2.7 Questions on explicit cueing and didactic memory⁹

The hypothesis that explicitness of instruction plays a role in pupils' attainment stems from two schools of research thought. The first—from psychology—highlights the role of cognitive representations and conceptualization in development and learning. The second—from sociology—focuses on the sources of inequality in performance at school and emphasizes the "collusion" between higher family socialization of students from higher-SES backgrounds and higher expectations of school achievement.

A.2.7.1 *The psychology angle: metacognition and cognitive clarity*

Over half a century of learning theory scholarship converges to underline the important role of what is often collapsed under the catch-all term *metacognition*: a set of self-knowledge and self-regulation processes that serve an ability to triangulate up and out from one's own mental processes and information being learned (Flavell, 1976). In the sphere of scholarship on the route into writing, Downing & Fijalkow (1984) conceptualized cognitive clarity as a conditional gateway to this route, borrowing heavily from foundational work by Fitts & Posner (1967). They lend central emphasis to cognitive clarity, defining it as comprehension of two types of reading-related concepts: the concepts tied to understanding the functions of the writing, and the more technical-centric concepts that we mobilize to talk about speaking and writing, to describe how they work.

A whole stream of research (Brossard, 1994; Brigaudiot, 2000) has since teased out and typed the metacognitive knowledge-sets that are mobilized in the route into writing and that link into the

⁹ This section borrows largely from the following article: Jacques Crinon, Natacha Espinosa, Marie-José Gremmo, Annette Jarlégan, Maria Kreza, Anne Leclaire-Halté. "*Cognitive clarity and teaching of reading and writing in the first grade: what teaching practice?*", *Pratiques* [ePub], 165-166 | 2015, published online 01 October 2015, accessed 25 November 2015. <http://pratiques.revues.org/2586>; DOI: 10.4000/pratiques.2586

various components of learning to read and write— knowledge-sets that span mental representations, conceptualizations (Cèbe, 2000; Chauveau, 1997) of the nature and functions of written language, the reading process, tasks in school.

Learning the alphabetic principle, or the principle of regular spelling patterns, as well as the goals and outcomes of tasks in school, opens possibilities to learn more, including, by analogy, grapheme-phoneme correspondences or spelling.

Are these conceptualizations explicit? Either way, as a rule, researchers can only grasp these conceptualizations through the way children verbalize them (see, for instance, the interviews on metagraphic experiences used to access orthographic representations—Jaffré, 2003). There is a large evidence base suggesting that one of the teacher's pivotal roles in scaffolding instruction is to bring pupils to this cognitive clarity, by explicitly teaching—and making sure students explicitly learn and clarify—how writing works and the strategies, aims and practices of meaning and challenge to them. The relationship between explicit and implicit knowledge also comes up as a wider question. Pupils, in reading as in writing, possess an array of implicit knowledge, knowledge that has been acquired through encounters with writing and texts without ever being taught nor even voiced out loud, but which experimental studies have evidenced are mobilized for structuring sequences of letters or for morphology (Gombert, 2004). However, “these skills embedded by implicit learning do not appear to be consciously accessible nor intentionally mobilizable by the reader” (Gombert, 2004, p. 50), which is why a lot of psychology researchers as so ardently behind explicit instruction of the alphabetic system (NRP, 2000).

Furthermore, being capable of connecting diverse learning-work situations and using earlier knowledge is shown to be a prerequisite condition for learning—note that failing pupils juxtapose classroom class situations without really connecting them up. This is where the importance of ‘guided connection strategies’ (Bucheton & Soulé, 2009) comes in, where the teacher explicitly articulates the connections and calls on “didactic memory”, a shared memory related to the learning objects taught (Brousseau & Centeno, 1991), especially when opening and wrapping up the classroom session. In the session opening phases, the teacher can employ a set of instructional processes of strategies that refer back to learning-work situations the pupils have already encountered or to earlier knowledge already constructed. This resituation gives pupils “the possibility to mobilize learning they did not fully possess, learning that they would not have used on their own but which enables them to give meaning to the task at hand” (Fluckiger & Mercier, 2002). To wrap up the session, the teacher can employ what Brousseau (1998) coined “institutionalization”. Indeed, without these moments for institutionalization, pupils’ progress will suffer (Coulange, 2011).

A.2.7.2 *The sociology angle: family socialization, school socialization, and misunderstandings*

Looking in through the lens of sociology adds the dimension of educational inequality, including on attainment of reading and writing skills, between pupils from contrasting socio-economic backgrounds. Readiness for the route into the written word entails a series of conceptualizations that go hand in hand with mental distance to language, the ability to think about language as object, and thus make the transition from a command of language that is purely practical up to a command that is “conscious, reflective and symbolic” (Bautier & Goigoux, 2004). However, as Lahire (2008) says, cognitive dispositions do not stop and start at ‘meta’ dispositions— they also encompass what, in differentiated and class-based social spheres, we can qualify as socio-political dispositions. Indeed, given that language practices are fundamentally intertwined with the relationships between

the agent-practitioners, the symbolic command of language, the capacity to adopt metalanguage dispositions, can—in certain social settings—equate to symbolic command over those whose command of language is purely practical. Compounding the issue, many children from low-SES homes come to school under-aware of the forms and expectations of schoolwork and therefore struggle to identify the learning targeted (Bonnéry, 2007).

Different ways the teacher teaches can either accentuate or attenuate these inequalities. The “invisible” factor of instruction (Bernstein, 1975/2007), reflected in the amount of implicit and uncertain, proves hugely differentiating, as it further clouds the landscape of understanding for pupils whose family socialization has not readied them for school. This is how “socio-cognitive misunderstandings” arise (Bautier & Rochex, 1997) between the teachers’ aims and certain pupils’ perceptions, and how “the gap widens between pupils effectively inside in requisite intellectual activity and those that are outside looking in, going through the motions as they only see the mechanics of the classroom task” (Rochex & Crinon, 2011, p. 12).

At this juncture, while English-language studies on the most effective teachers find that they explicitly articulate their instructional approaches and procedures, even going as far as leading pupils to appropriate a “frame of instruction”, i.e. understand the objectives of school (Carette, 2008), the research we have just explored prompts us to think that this same practice could equally characterize the most equitable teachers, i.e. those that look to take socially-acquired dispositions out of the equation and help children from low-SES homes to build at school what the other children have often already built at home.

A.2.7.3 *Research question*

We built up on this theorization to formulate the general hypothesis that explicit instruction—in myriad forms—can play a positive role in educational attainment, especially for struggling pupils and pupils who are less socially/culturally literate in classroom practice.

We will see whether that hypothesis is verified.

A.2.8 *Questions on class climate and pupil engagement*

The aim of this line of research is to investigate how class climate and pupil engagement are linked to reading and writing performances.

A.2.8.1 *Climate–engagement–performance linkage*

A positive classroom climate can be hallmarked by the quality of in-class relationships, both teacher–pupil relationship and pupil-pupil relationships, and by the enthusiasm that the teacher injects and the pupils share. Several studies led in Anglosphere countries find that these dimensions, whether taken individually or separately, are predictors of pupils’ reading performances in the last year of preschool and in first grade (NICHD ECCRN, 2003; Pianta, 2003). Here we set out to determine how far the influence of class climate on pupils’ performances in the first-grade classes of our sample adds to the influence of the teachers’ instructional interventions. In other words, we aim to establish how much better first-grade pupils’ progress in a positive class climate.

Furthermore, a class offers more learning opportunities when the pupils are behaving appropriately, constantly working, and effectively engaged on-task (Blair, 2003; Raver, 2004). We

will therefore also examine the effect-size of these classroom organization-related variables on pupils' learning outcomes.

Data on class climate and pupil engagement was collected and compiled to determine how far these variables shape reading and writing education at first grade in France.

A.2.8.2 *Construct a control variable*

Study into “first-grade reading and writing education” is all about measuring the opportunities for learning provided to a class via the metric of time given to instruction on the various components of reading and writing.

Nevertheless, allocating time to instruction on a given skill will obviously be a meaningless exercise if the pupils are not engaged in the task or if a negative classroom climate is a barrier to learning productively. Our analytical strategy thus involves constructing “climate” and “engagement” variables as control for teachers' instructional styles and strategies, which are the central variables of this research write-up report.

A.2.8.3 *Adapt and validate a measure of class climate and pupil engagement*

Our rationale for integrating instructional dimensions into our general explanatory model was not just sound science but also to validate an American observational instrument for assessing classroom climate and engagement, the CLASS (Classroom Assessment Scoring System) devised by Pianta, La Paro & Hamre (2008) (see A.3.4.1.7). This instrument holds huge methodological value for educational research, but despite widespread adoption for assessing effectiveness of in-class practices in North America, it has never, to our knowledge, been used in France.

Our ongoing multilevel analysis will serve to distinguish the most effective classrooms from the least effective classrooms, after which we will move forward by using CLASS to evaluate the 15 most efficient and 15 least efficient classrooms based on video-recorded observational evidence. The expectation is that this research effort will validate the use of CLASS in French-language and in the French school system, as already done for Spanish in Chile and Finnish in Finland (Leyva *et al.*, 2015; Pakarinen *et al.*, 2010), which should then open up broad perspectives for further educational science research.

Within the framework of research reported here, we assessed classroom climate and pupil engagement using a downscaled number of items loosely adapted from CLASS. Our inquiry was thus set up to find out how far this shorter purpose-adapted version manages to capture classroom climate and pupil engagement by comparing the data it obtained against the data obtained using the full-version instrument.

In short, our goal is to determine whether classroom climate and degree of pupil engagement on programme tasks have an effect on pupils' attainment of reading and writing.

A.2.9 *Questions on differentiated instruction and support to struggling pupils*

The overarching objective of this research is focused on the effectiveness of practices for teaching reading and writing in first grade, but a specific sub-objective is focused on the effectiveness of these practices for pupils most at risk of school failure.

Effort to investigate the practices that work best for progressing the pupils starting out first grade the furthest behind on performance means exploring at least three dimensions that, from an institutional standpoint, dictate the options for supportive action to struggling pupils:

- 1 Do the regular teaching practices addressing the whole class have different effects depending on pupils' initial level? Are some practices better geared for the pupils furthest behind than for the others?
- 2 Do the regular teaching practices tailored to pupils most at risk of school failure have positive effects on learning outcomes?
- 3 Do out-of-class practices delivered specifically to pupils most at risk of school failure have effects?

For the first sphere of action, there is ample scientific literature to show that, on balance, a certain number of practices addressing the classroom as a group actually have more effect on the pupils initially furthest behind (Bressoux, 1994; Bissonnette, Richard & Gauthier, 2005). One of the objectives of our research is to confirm and consolidate whether certain practices studied have this kind of effect, so there is no need to expand on this issue here.

For the second sphere of action catering to struggling pupils, the practices that teachers can tailor to pupils most at risk of school failure are often collapsed under the somewhat hazy term of "differentiated instruction". However, the range of practices qualifying as differentiated instruction is so loose that it would not be a relevant exercise to inquire into its effects here. Moreover, Jobin & Gauthier (2008) showed that research on differentiated instruction fails to demonstrate substantial strength of evidence of effectiveness. That said, a recent review of the literature attempting to precisely chart the ways teachers accommodate struggling pupils did highlight a number of practices that can be grossly collapsed under one of two macro-strategies (Descampes *et al.*, 2008; Piquée, 2010). One, teachers schedule special tasks for their struggling pupils into their lesson plan. The task content is differentiated, i.e. different to that delivered to the rest of the class peer-group, but on the premise that it is adapted to the needs of pupils fallen furthest behind. Two, teachers do not schedule any special content, and all pupils get the same tasks to do, but the process is differentiated (extra time, help with the method, small-group tutorials with closer intervention from the teacher, and so on). Thus second form of support is thought to better serve the pupils' progress (Bressoux, 1994; Duru-Bellat, 1996; Descampes *et al.*, 2008; Piquée, 2010) as it counters the processes of pigeonholing and simplified assignments, but the evidence base remains scant and the hypothesis warrants more expansive investigation.

For the second sphere of out-of-class action, a number of international studies demonstrate the effectiveness of compensatory one-to-one education, often delivered daily through ready-programmed protocols (D'Agostino & Murphy, 2004; Elbaum, Vaughn, Hughes & Moody, 2000; Swanson, 1999; Viriot-Goeldel, 2011; Suchaut, Bougnères & Bouguen, 2014). In France, where protocols like this is uncommon, systems of support to struggling pupils have posted a mixed bag of results (Piquée, 2005; Suchaut, 2009). Some systems demonstrate relative effectiveness (Piquée, 2003; Piquée & Suchaut, 2004), while others are found to have zero effectiveness or even counterproductive effects (Mingat & Richard, 1991; Piquée 2007; Goux, Gurgand, Maurin & Bouguen, 2013). To further our grasp of this third sphere of support action for struggling pupils, we will narrow our focus exclusively on support structured through schooling under national education council stewardship.

In France, over and above the support teachers provide their pupils inside the class, struggling students can get further out-of-class support through extra instruction qualified as complementary tutoring (acronymed 'APC' in French). The 'APC' scheme brought in at the start of the 2013 school year provisions a total of 36 hours of instruction per year per class "to help pupils who are experiencing difficulties learning, to support them in their individual education, or include them in any other activity scheduled under the school's education action programme and—where appropriate—the local-authority project for education action" (MENESR [Ministry for Primary, Secondary and Higher Education and Research], 2013).

As "the provision of support from the schoolteacher [...] may not be enough for some pupils" (MENESR, 2014), special education is made available through French special education networks for struggling students, acronymed 'RASED'. Each RASED can mobilize CAPA-SH-E educators—CAPA-SH is vocational training for special needs education in primary school, and *E* is specialism *E*, i.e. inclusion through instruction—whose mission is to provide "special needs support to pupils with learning and comprehension deficits". The specialized educator's mission is to liaise with the class teacher and/or education-stage teaching staff to prevent or resolve learning difficulties for these children and provide appropriate early remedial instruction within the framework of a special needs programme" (MENESR, 2014). CAPA-SH-G educators—where *G* is specialism in play therapy—are assigned missions defined in these words: "Special needs educators specialized in play therapy provide special needs support to pupils for which situational analysis shows a need for guided direction to help meet the school's codes and requirements and foster or restore appropriate investment in classroom activities. The specialized educator's mission is to liaise with the class teacher and/or education-stage teaching staff to prevent or resolve behavioural difficulties for these children and implement actions to foster progression, within the framework of a special needs programme" (MENESR, 2014).

To examine the effects of these special out-of-class (out of teachers' classroom hours) action schemes, we would need to design and deploy a whole different assessment protocol, not least because individual pupils can potentially cumulate these aids and services, which would makes it difficult to unravel the effects and effect-sizes when these system accommodations are layered up. Consequently, for this third sphere of support action, the objective will doubtless be to focus foremost on charting how intensively this form of support gets used. Then, and only then, could we start thinking about whether or not to correlate the use of these interventional support schemes with pupils' progressions.

A.3 Methodology and analytical strategy

A.3.1 Identify the effective practices

This research is not a trial aiming to empirically verify “that what we think is the right instructional principle actually works effectively in practice” (Dehaene, 2011, p. 102): we are not setting out to verify that a method we think relevant produces positive effects. That said, we are setting out to identify the hallmarks of the mainstream-curriculum instructional practices that prove most effective. We thus intend to move forward with research-based education studies that have assessed the size of the ‘teacher effect’ in elementary school (see Attali & Bressoux, 2002 or Bressoux, 2011 for a review): we are looking to identify the components of instructional practices that help close inequality gaps in pupils’ learning.

On the didactics front, our objective is to identify which contents, tasks and instructional strategies work best for pupils’ reading and writing education at different periods of the school year. We want to find out which trajectories of pace, depth and mix were differentiated for different pupil audiences targeted. We also want to resolve the issue of whether the content defined in statutory curriculum programmes should be taught sequentially or simultaneously, isolatedly or integratedly, explicitly or by straight reiteration. To illustrate, do the teachers that devote the bulk of their teaching time to decoding get better all-round results than teachers whose policy is to focus instructional design around text comprehension skills or teachers who mix reading and writing instruction side by side? Another question we want to address is whether practices that prove effective share common traits, and if so, which, without any preconception that such traits are linked to textbook factors. Finally, we also aim to resolve the issue of whether the factors of effectiveness identified in experimental or interventional research translate up to natural context, i.e. in contexts not transformed by the research frame itself.

Working within a ‘wide net’ paradigm, we designed a system of inquiry that is built around field observations on a very large sample of pupils and teachers at work in an array of school contexts and endowed with the scientific method needed to demonstrate causality in social science (Courgeau, 2000; Dumay & Dupriez, 2009).

The first requirement consists in capturing the effects of practices in terms of educational “value-added”, which means separating pre-study from in-study-period learnings. It is only the pupils’ learnings in first grade that has been matched up against the teachers’ practices. To isolate these learnings and neutralize effects spilling over from before starting first grade, we measure pupils’ competences using pretests at the start of first grade and posttests at the end of the first-grade school year followed by delayed posttests at the end of second grade to find out whether pupils’ progressions are variable, sizeable and lasting.

A second requirement in our analytical strategy consists in capturing the effects of practices “all other things being equal”, or put more precisely here, “all other variables included in the model being equal” (Bressoux, 2007, p.74). This is because on top of instructional factors, there are numerous other factors that can shape and be statistically linked to pupils’ growth. Typical examples would be pupils’ sociodemographic and schooling factors, structural classroom factors, and other

factors that we have inventoried in detail further down¹⁰. To be able to separate out the effects of various different variables liable to affect pupils' progression and tease out their respective impacts, we employ multilevel models which serve to estimate the impact assignable to each variable independently of impact of the other variables included in the model (Bressoux, 2010).

The third requirement is tied to the approach to describing teachers' practices. The main, and widely signalled methodological weakness of studies framed in the classroom ecology paradigm stems from the operationalization of teacher-level action variables, which all too often relies on teachers' own declarations collected through self-report questionnaires as it is difficult to study a vast cohort through any other means (Fijalkow & Fijalkow, 1994; Sensevy, 2007). However, the validity of the results obtained is weak, as declarative practices have to be taken as actual practices. This makes direct observation the only reliable way forward (Arnoux, Bressoux & Lima, 2008). Direct observation applied on big classroom samples provides a quantitative backbone that makes it possible to statistically control for certain variables and demonstrate causality (Jarlégan et al., 2010). However, it rarely gets employed in practice, as it requires raising an army of field investigators. Here, by reaching out to teams from 13 university faculty laboratories, we managed to raise this army.

The fourth requirement is to study mainstream-curriculum classes in everyday situation. To meet this requirement, we reached out to teachers in fourteen different *académies* [regional education authorities], asking them to simply keep teaching as they normally would. The very first contact was couched in these terms: "Our objective is not to demonstrate the superiority of any one method, but simply to identify—with no preconceptions whatsoever—the characteristics of the practices that prove most effective and most equitable. For this purpose, we are looking for experienced primary school teachers who are teaching to mainstream curriculum but interested in finding out more about what actually shapes the quality of their pupils' learning. We would therefore be delighted if you would consent to participate in our study by letting one of our field-researchers—who would be the same person all year through—come in and talk with you, analyze your pupils' progressions, and observe your teaching practice". A host of teachers answered our call for participants and opted in—some seeking confirmation of the didactic and instructional theory that they staunchly defended, while others—the majority—simply wanted to know whether the actions they adopted actually work.

Setting ourselves **a fifth requirement**—to construct a population sample rich with variety—we enrolled experienced teachers working in a diverse range of geographic zones and social-background contexts. The population of pupils ($n=2507$) counts 50.3% girls. Social make-up is a varied mix, slightly under the national-average SES (see Annexe 1), and 4.2% of the pupils have fallen behind in school. At class level ($n=131$), 27.4% are in a priority-education setting and average class size is 22.2 pupils (versus 16.7% and 22.7 for the year-2013 national estimates; MENESR, 2014). At teacher level, average in-tenure service is 17.4 years (SD: 7.4), of which 8.4 years in first grade (SD: 5.3). The teachers enrolled appear to work to contrasting sets of didactic and instructional principles. To illustrate, 91 (i.e. 69% of the sample) used textbook readers while 40 did not. Among the 91 textbook-users, 13 (i.e. 10% of the sample) had opted for a synthetic phonics

¹⁰ This analytical strategy approach enables us to conform that the observed relationship between target factor and pupils' final performance outcomes is not driven by another variable, which may itself be linked to or partially confounded with the factor under study. Put another way, the net effect of the practices gets ironed out as we rigorously control for pupil factors and initial performances together with the class factors.

approach, 15 a pure phonics approach, 2 a quasi-whole-language approach, and 61 an embedded approach (Goigoux, 2003)—a breakdown that reflects much the same diversity as described in a recent large-scale telephone interviewing survey (Deauviau, 2013).

A.3.2 The analytical model

Instructional practices are the target variables of this study, for which we want to measure effect-size. To isolate the effect—contributions of these practices, we have to consider and neutralize the effects of the other factors affecting pupils' performances. These other factors operate on two order-levels: some are individual variables characterizing the pupils, while the others are context variables tied to schooling that either grows the pupils or holds them back (class-factor variables, teacher-factor variables, etc.). Our analytical model (see Figure 1) schematizes various linkages that should be controlled for.

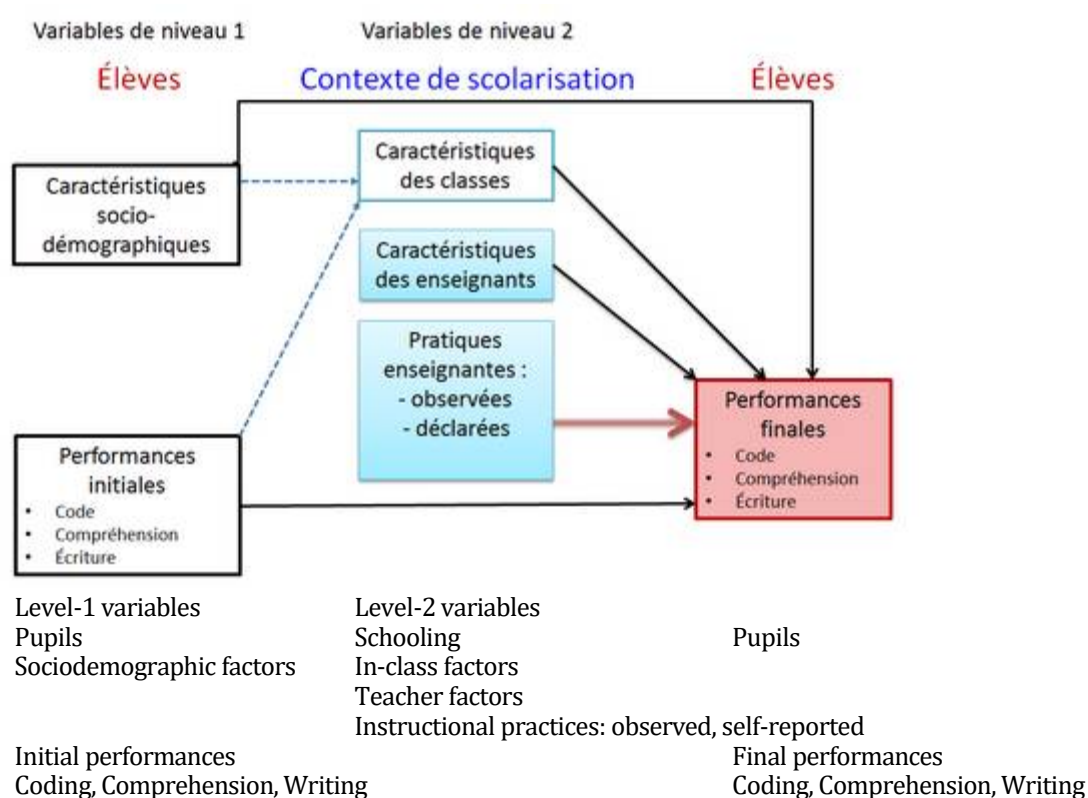


Figure 1: The analytical model

Key: $A \rightarrow B$ means that B is at least in part determined by A.

The double arrow (in red) originates from the target variables.

The thick-line black arrows originate from control variables.

The dashed-line arrows indicate the source of the aggregate variables.

Statistical analyses will attempt to isolate the contributions of the teaching practices to pretest-posttest differences in pupils' end-of-year test scores after controlling for the impact of other factors. We will then evaluate the effect size of these practices by analyzing the relative position of each pupil in relation to the hierarchized positions of the all-pupil superset.

The data collected and reported below is hierarchical in that it comes through two levels of observation and analysis, i.e. pupil level (level 1) and class level (level 2), linked by a nested relation¹¹.

Level 1 encompasses the pupil-factor variables. As is now typical in the educational effectiveness research scholarship, we control key variables:

- child's age and gender,
- whether or not the child has already repeated grade,
- child's father's and mother's socio-economic status attributes (income-occupation–education) as per the INSEE [French national statistics bureau] classification scheme
- the language(s) spoken at home.

We add to this first level the results of assessment on the pupils' initial levels of competence, based on three key domains of ability to read and write: oral and written code (e.g. phonological competence and names of letters), comprehension of texts (read aloud by the adult at year-start), and writing. To obtain a reliable dataset, all the competence assessment tasks that we either borrowed or designed ourselves were administered individually or in very small groups in the first week of September and then in the last fortnight of June, by the research investigators and a field squad of a hundred-odd pre-trained Masters students who were paid for their work. The standardized tests administered to the children are independent of the instruction assessed. Although some were original, most were borrowed from earlier experimental studies, including the study supervised by Gentaz (2013), so as to facilitate cross-connection and comparison between findings from research led in different paradigms. Given the tight framework of this synthesis report, we elect not to detail the variables characterizing pupils' performances so as to keep a clean focus on the more original variables that we constructed here to capture the effects of teachers' practices¹².

Level 2 is the instructional context of each class. This instructional context is described using four kinds of variables characterising:

- the class-groups, captured by aggregating the individual pupil characteristics of each class: social mix of the class (computed from the socio-economic status attributes of the parents), level of the class (computed from the pupils' mean initial performance levels), heterogeneity of the class (computed from the standard deviation on the pupils' initial performance levels).
- the pupils' environment: geographic location, priority-education setting, class size (recorded each 3-month term), number of teachers, existence of any special accommodations ('CLIS' [mainstreaming classes], RASED or other aides and assistants), schoolweek schedule (8 or 9 half-days), and other relevant data.
- the teachers: age and gender, initial, university and pre-service/in-service teacher education, experience in the job (years in teaching and years in first-grade teaching).
- the teachers' actual practices.

¹¹ We could have used more levels in our model, given that the pupils (level 1) are grouped in classes (level 2) nested in schools (level 3) governed by constituencies (level 4) coordinated in *départements* (level 5), and so on. However, with no hypothesis on higher levels to go on, we opted to stick with the 2-level model.

¹² All the information collected has been anonymized. Each researcher has access to the full dataset, made available via a secure website, as per the principle of data co-ownership and co-sharing that our alliance signed up to in February 2012. The lead project manager, engineer and statistician are the only people in possession of the codes needed for re-identification of classes. However, each field-investigator has been given the code number of their class they followed,

The teachers' practices are the target variables of this study and were constructed from a small selection of indicators that prior research had shortlisted as good candidates for explaining effectiveness differentials between contrasted teaching practices (Nonnon & Goigoux, 2007). The selection pools two kinds of indicators. One is a set of didactic strategy indicators, which are specific to the key instruction as defined in the French elementary school education programmes (MEN, 2008) or as emerged from the findings of didactic research in French, in cognitive psychology and in language science: instruction on the alphabetic system and word identification processes, instruction in reading and listening comprehension, instruction in producing writing. The other is a subset of structural indicators that educational science has shown to be important: explicitness of the teaching connected to guided mobilization of the class's didactic memory (Rosenshine, 1986; Rochex & Crinon, 2011); the formats and frequencies of differentiated instruction (Piquée, 2010); classroom climate and pupil engagement in tasks (NICHHD, 2002). The next section brings an in-depth presentation of these indicators.

B.3.2 Observe the teachers' practices

We elected to observe the teachers' practices from three angles, in succession, as set out below:

- we study the reading/writing teaching offer by examining the type and time of tasks given to pupils and recording the traces of teachers' and pupils' activity;
- We analyze the programme of study adopted for teaching grapheme-phoneme code in tandem with the text and textbook resource materials that the teachers use;
- we also map the conditions that influence quality of learning by assessing the explicitness of the teaching delivered, the strategies of differentiation instruction employed, the pupil grouping systems implemented, plus classroom climate and pupil engagement.

A.3.3.1 Analyze the teaching offer proposed to pupils

A teacher's biggest resource—and consequently their biggest constraint—is the time that they have for delivering instruction to their pupils. The teacher has a statutory teaching time for French, set by education law authorities at 10 hours per week in first grade¹³, and so a number of strategic decisions to make in order to organize their education offer. To achieve our research objectives, we have purpose-devised a set of tools to capture and translate these decisions, i.e. to identify the programme-of-study content, the programme-of-study schedule, the tasks given to the pupils and the way the teachers teach. Put another way, we built tools describing the learning opportunities that each teacher makes possible through the tasks they propose to their pupils.

For this purpose, we adopted a unit of observation of three full weeks of reading–writing instruction, with one unit surveyed per 3-month term. Each field-investigator thus observed and filmed thirty-odd hours of French language sessions in November, then again in March and once again in May. This temporal sampling framework was expected to capture the hallmark characteristics of primary-school teachers' mainstream instructional practices, as each teacher had agreed simply to keep

¹³ Teachers have the flexibility to introduce content earlier or later than set out in the programme of study, as long as they deliver the annual teaching hours stipulated for each subject area on curriculum.
<http://www.education.gouv.fr/bo/2008/hs3/MENE0813208A.htm>

teaching as they normally would and to keep us informed of any one-off events coming up¹⁴ (such as a school trip).

The observations were conducted simultaneously in and across all the classes by 58 tenure-track lecturers who had been members of the research group from the outset in 2011, 27 PhD-student/PhD-graduate teacher trainers, and 35 in-service educational advisors and another 18 retired from national education service (former educational advisors or primary-teacher trainers, volunteers). All these field-investigators were pre-trained, worked to a 60-page-long statement of field work, and posted the output from their observations at the end of each week via a purpose-designed protected-access website. Their job was to record, stopwatch in hand, the path and process of all the in-class sessions, while a fixed camera set in $\frac{3}{4}$ -rear shot position was recording the sessions onto film, thus keeping a secure trace of these sessions to fall back on¹⁵.

We now have nearly 3000 hours of video evidence stored at the IFé *LireEcrireCP* site at ENS Lyon, and which all the researchers in our project group have access to. The expectation is that this access will help the team to qualitatively examine patterns and phenomena not shortlisted for the quantitative analysis strand of the survey, and thus refine the interpretation of results and findings obtained through statistical processing. The most and least efficient and/or equitable classrooms, for instance, will serve for complementary analyses.

A.3.3.2 Study the tasks and teaching time allocated

Our objective is to correlate the learning opportunities proposed by the teachers with the learning performance outcomes of the pupils. At this point, here, as with any research of this type, the theorization and methodology problems to resolve will revolve around identifying these learning opportunities and choosing the allied grain of observation and analysis. The strategy chosen was to measure time allocations, and to do it at a relatively mid-sized grain of analysis¹⁶, i.e. the tasks assigned to the pupils. The classroom sessions observed were thus broken down into temporal units characterizing the work assigned by the teacher. In other words, we describe what the teacher asks his pupils to do.

We define a task¹⁷ as a function of the goal the teacher wants their pupils to attain under given conditions (Goigoux, 2002). The field-investigator identifies a switch of task as each time the teacher gives the pupils a new goal or changes the conditions for attaining a goal. Clearly, then, the focus of our descriptive work is not the activity behaviour of each pupil but the activity behaviour that the teacher expects to see— an approach that enables us to recompile the time allocation and temporal sequencing of tasks given to the pupils, which we have dubbed as their weekly ‘time-budget’. In short, we are attempting to precisely qualify what the pupils are supposed to be doing, in response to more or less explicit direction from the teacher.

To get to a position where we were able to homogenize the temporal breakdowns of in-class sessions and, from there, assign each task to a pre-established category, we produced a tasks

¹⁴ If so, the day scheduled was replaced by rollover to another consecutive day.

¹⁵ The pupils parents’ gave consent to these shots on the caveat that we promised not to disclose the films publically and not to use them for any purpose other than coding the teaching sessions.

¹⁶ On a gradient where analyzing teacher-pupil interactions would make a finer grain while analyzing how the week is broken down into sessions would make a courser grain.

¹⁷ Our reference frame here is the construct of task as applied in human factors and ergonomics, a definition that “carries with it the idea of prescription, if not obligation, whereas the construct of activity is framed in terms of processes mobilized by the subject to execute prescriptions, fulfil obligations” (Leplat & Hoc, 1983, p. 50).

typology model, which can be found in Annexe 2. It took a whole year of hard effort to hammer out a scientific consensus between the contributing researchers and then test that our typology held validity (verify that all the tasks used in first grade were classable), reliability (interjudge agreement), and operable manageability (reasonable range of categories, prototypical exemplars). In a coherent fit with our research questions, our typology model is structured into thirty-one types of tasks, which roll up into five big supersets corresponding to the big subject areas for teaching to read and write in first grade: phonological awareness, reading, comprehension, writing, study of language. Thirteen tasks require learners to also decode the kind of linguistic unit they mobilize (letter, syllable, word, sentence or text), so as to distinguish, for instance, between spoken syllables and spoken sentences.

This typology ultimately discerns seventy-three different types of tasks, enabling analysis of the teaching offer to drill down through a progressively finer-grained gradient of four levels:

- global time-budget (minus time-off-task)
- sets of tasks (5 categories),
- tasks (31 categories),
- tasks per type of linguistic unit (73 categories).

In order to study the way teachers operationalize explicit teaching, we created a sixth bucket, headed “didactic memory” (DM), to isolate tasks directed towards the institutionalization of learning skills (DM2) or the explicit recall of pre-scaffolded knowledge (DM1). The field-investigator records which type of content these tasks address, so that they can later be stratified under the above categories when computing the time-budgets.

Our typology also features a bucket headed “Other” to compile any unclassable tasks, and last but not least, a very important bucket called “Time off task” to capture episodes lasting over a minute that are disengaged from reading or writing instruction. This would typically be a straight session suspension (to roll-call pupils scheduled to eat at the canteen), slack time waiting when the pupils have finished their taskwork, downtime to relax, time-outs for discipline and classroom control, a maths exercise given to a small sub-group while the rest are reading with the teacher, handing out resource material, cutting out stickers or gluing photocopies, and so on. With this methodology, after subtracting the time off task, we get a sharp and accurate measure of actual teaching time available for reading/writing instruction.

A.3.3.3 Task coding system and technique

To recompile the time-budget of a class, each field-investigator attends all the reading–writing sessions provided to the pupils during the three test weeks. In observation, they only code the tasks given to the first-grade pupils, including in essentially non-French content areas sessions if the teacher had previously given a heads-up that the lesson plan was to also include a reading or writing activity.

The field-investigator works in two timeframes. The first timeframe is ‘live’, where they take notes on the spot, as and when the sessions unfold, recording the events observed in the order they occur. They write down what the teacher says to the pupils or else a summary of what the pupils are asked to do. They note in the margin the start-time of each new task, following a stopwatch set running at

the start of the session. The time-breakdown is thus recorded ‘one the fly’ each time the field-investigator identifies a task switch¹⁸. Under another header, they add information on the explicitness of the teaching and on instructional differentiation or classwork modalities. If the class is split into groups of pupils on different tasks, then the field-investigator records the tasks given and the number of pupils engaged in work without the teacher, and later takes photos of the teaching material handed out.

The second timeframe is later but ‘almost live’ (generally at the end of each half-day or day), where they encode all the information gathered and key the codes into a pre-formatted Excel spreadsheet featuring a new row for each new task (see the illustrative example below).

A	B	C		D	E	F	G	H	I	J	K		L	M	N	
N° séance	Chronomètre	Tâche 1 avec M			Explicite	Support	Groupement	Type d'écrit	Différenciation	Tâche 2 (sans M)		Tâche 3 (autre M)		Commentaire		
Ordre dans la semaine	Noter le début de la tâche	Quelques mots clés	Code T1		EP - EC - E	F, E, A, C, M, L, D, O, T, TB1, AUTR	C - I - PG - GGPG - AUTRG	NT - AFF-CR-DOC-LEG-LIS-PO-REC-TTE-MES-AUTR	TIA - TD - TDA-TB - DIFAUTR	Nb d'élèves sans M	Code T2		Nb d'élèves avec autre M	Code T3		
5	0:52:30	temps mort														
5	0:53:00	dictée graphèmes	E4-L			C	GGPG	NT	TIA	18	PG1&PG5&PG4				aide à la fusion	
5	0:58:00	dictée syllabes	E4-S			C/T	GGPG	NT	TIA	18	PG1&PG5&PG4				M encourage, aide à la segmentation, à la fusion, renvoie à l'écrit	
5	1:02:00	temps mort													M ramène le calme au sein du grand groupe	
5	1:02:30	consigne lue en autonomie	PG5&PG4			F	GGPG	NT		18	PG1&PG5&PG4					

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Sessio n #	Stopw.	Task 1 w/ T		Explicit	Support	Grps	Type of writing	Diff Inst	Task 2 w/o T		Task 3 Other T		Comment
	Record task start- time	Few keywords	Cod e T1	EP-EC- E	F, E, A, C, M, L, D, O, T, TB1, AUTR	C-I-PG- GGPG- AUTRG	NF-AFF- CR-DOC- LEG-LIS- PO-REC- TTE-MES- AUTR		Nbr w/o T	Code T2	Nbr w/ other T	Cod e T3	
5	0:52:30	timeout											
5	0:53:00	grapheme dictation	E4-L		C	GGPG	NT	TIA	18	PG1&PG5& PG4			helps with blending
5	0:58:00	syllable dictation	E4-S		C/T	GGPG	NT	TIA	18	PG1&PG5& PG4			T encourages; helps with segmentation, blending, connects up to class-wall materials
5	1:02:00	timeout											T reinstills calm in the big group
5	1:02:30	guidance on self- worksheet	PG5 /PG 4		F	GGPG	NT		18	PG1&PG5& PG4			

Table 3: Excerpt from a “Tasks” spreadsheet—10-minute segment taken in week 21, on class 102

Time unit of survey is the week: all the in-week information collected is gathered into a single Excel file. Each week is broken down into class sessions (column A) which are in turn subdivided into tasks (in rows in the table). Each session conducted gets numbered in chronological order over the course of the week, and is inventoried in a file called “Timetable” so as to index the videos simultaneously recorded.

Tabulating all the tasks that make up the in-week sessions in a continuous sequence makes it possible to run an automatic computation of the time-budgets, i.e. the mean teaching time provided to a pupil for each type of task. When not all pupils are doing the same thing, the timeframe is weighted by the headcount of pupils who are on that task. The task can be practiced with the teacher present (columns C and D) or with pupils working on their own (columns J and K), or possibly with another aide/assistant (columns L and M): contract teacher, special needs educator, or any other adult present in the classroom.

If a teacher-directed instruction given to the whole class appears to embrace several types of task, the field-investigator chooses the dominant task and records just one code (in column D) for each task observed. Tasks lasting less than a one-minute interval are not encoded. There are other

¹⁸ A switch in task is a change in task goal or a change where pupils are asked to achieve same task goal but under different conditions.

equally short tasks that come up on many occasions, entangled with other tasks, in which case the field-investigator creates just one row, for the major task, but records a code pair using a slash (e.g. C4 / SoL1). Our computerized data capture and processing system¹⁹ then assigns $\frac{2}{3}$ of the time to the first task recorded and $\frac{1}{3}$ to the second.

Moving along to column E, the field-investigator records whether the teacher gives or asks for, at least once for that task, an explicit explanation of the “Why” (when there is explicitation needed of the outcomes of the task) or the “How” (when there is explicitation needed of the procedures and strategies or the knowledge mobilized or to mobilize for getting the task done).

Moving along to column F, the field-investigator states the support materials used on the task: sheets or photocopies (individual), restickables (individual), slate, classbook or flyleaf, textbook, storybook, featurebook or other early readers, computer, tablet, blackboard or whiteboard, interactive whiteboard or other classroom displays.

Moving along to column G, the field-investigator states the instructional teacher-led configuration chosen: all-class teacher-led (whole class-group receiving instruction, with interactions), individual (all children working separately, without interaction with peers), several small groups (with interaction between pupils, organized into ‘tables’: more than two groups of 2-plus pupils), on big group + one small group (with the teacher leading one of the two), or other.

Moving along to column H, the field-investigator code-records the type of writing required for the task: a poster, a report of an excursion, film or experiment, a documentary or explanatory text, a key (to pictures or diagrams), a list (menu, ingredients, etc.), a poem, a nursery rhyme, a song, a story (or excerpt from a story), a recipe, a message, a letter, or other. The field-investigator stratifies non-text, random sequences of syllables, words or sentences, pretend reading or writing.

Moving along to the last column I, the field-investigator states any differentiated instruction they are able to pick out: aids provided on identical tasks, different tasks (or accommodations) for certain struggling pupils, aids provided on different tasks (or accommodations) for certain struggling pupils; compacting and enrichments, i.e. different tasks for the highest achievers; or other.

A.3.3.4 Collecting tracks of teachers’ and pupils’ activity

There is no room here to fully list all the information collected by the field-investigators, so we will look here at two that give a glimpse of the kind of analysis made possible by this inquiry.

- Tracks of the pupils’ writing-task activities: each field-investigator was to photograph everything that two pupils—anonymized pupil ‘A’ and pupil ‘B’—had produced over the course of the test weeks and upload the photos to the research group’s collaborative platform. To date, we have around 6000 digitally-compiles pages. These two pupils had been hand-selected by the teacher in November: pupil ‘A’ was a “strong” pupil, well engaged in learning to write (already a reader and/or writer), while pupil ‘B’ was a “struggling” pupil, i.e. a pupil the teacher had singled out as worryingly at risk of under-attaining (on reading and writing).
- Tracks of the metalanguage terms used in-class: each field-investigator was to record the grammatical metaterms that they found evidence of on classroom-wall materials or in the

¹⁹ Yvonnick Fesselier is the IfÉ–ENS Lyon engineer responsible for designing and developing the whole computerization strand of our research project, from the website through to the data extraction system. Without him, none of all this would have been possible.

French-language classbooks, readers or files used by the pupils during the week (e.g. 'sentence', 'sight words', 'plural' and/or 'singular').

A.3.3.5 Recompile the programme-of-study schedule for code instruction

Another of our research questions addressed the quality of the different formats of programme-of-study scheduling, such as the programme of study of grapheme-phoneme correspondences (GPCs), as it articulates scientific dialogue with advocates of the idea that early, explicit, systematic teaching of decoding facilitates initial learning of reading–writing (Sprenger-Charolles & Colé, 2006, in the wake of the National Reading Panel Report; NRP, 2000). Like them, we take the stance that pupils will make better learning progress if they get the earliest opportunity to acquire skills that spark curiosity for how the alphabetic system works and facilitate self-learning while at the same time cutting short room for discouragement and groping in the dark (Goigoux, 2000). This is a stance that has been defended for nigh-on two decades (Gombert *et al.*, 2000), reinforced by the conclusions of the 2003 consensus conference (PIREF, 2003), but that has never actually been put to test in a French-speaking country. This was our rationale for setting out to confirm whether this position is effectively valid and, more importantly, under which contextual conditions early and systematic code instruction actually proves sizeably beneficial. Specifically, we wanted to know whether all pupils draw the equal benefit from code instruction programmes according to their initial level and whether *tempo* of curriculum delivery influences quality of learning directly or else in interaction with other didactics variables (writing, writing as culture of learning, etc.) (Bucheton & Soulé, 2009; McGill-Franzen, 2010).

On the methodology front, our investigations here were bracketed to just the first ten weeks of the school year, as the teachers' initial strategy choices heavily dictate their subsequent activity and heavily and lastingly influence pupils' quality of learning (INSERM, 2007). We inventoried all the grapheme-phoneme correspondences explicitly taught by the teachers over the first few weeks of the year (surveying their written tracks in pupils' classbooks, in schools' textbooks, in classrooms' posters and wallcharts, in teachers' lesson plans, etc.). To homogenize the standard of collection, we devised a matrix of the 63 main grapheme-phoneme correspondence pairs in the French language (an (am) ↔ [ã] is one digraph pair, for example) and 12 silent letters (like the 'p' in the French word *loup*). The field-investigator also flagged the path taken into the correspondence, i.e. the reading path, from grapheme to phoneme, or the writing path, from phoneme to grapheme. This technique accounts for all didactic pedagogies, whether phoneme-first or grapheme-first (Krick, Reichstadt & Terrail, 2007).

This inventory effort conducted in 131 classes leads to describe three core hallmarks of different programmes of study:

- their *tempo* (how many grapheme-phoneme correspondences were studied in a 9-week window);
- their theoretical return (using a frequency table for each GPC pair, we compute what percentage of a standard text is decodable at the start of week 10 of class);
- their actual return (what percentage of the texts given to pupils in week 10 is actually decodable).

This same analysis also serves to compare what the teachers actually teach against what the textbook authors advocate.

A.3.3.6 Inventory the textbooks

To test for any textbook effect and inform debate on how didactic resource tools influence teaching work, we inventoried the textbooks used and sorted them into categories in order to create a categorical variable that would work in the statistical model. The objective is to identify the relations between choice of textbook, programme-of-study schedule for code instruction, and the teachers' actual practices²⁰. Special focus is given to differences in practices between teachers using the same textbook, and overlap in practices between teachers using different textbooks.

Here, we collapsed the textbooks into four categories based on whether the textbook authors' approach is:

- embedded phonics: the authors prone combining explicit, systematic grapheme-phoneme correspondence instruction with writing and producing text, comprehension of text, and enculturation. Their approach is scaffolded in storybook readers, which may or may not be textbook-integrated, sometimes cut short and often broken down into episodes. The authors aim to construct an instruction-facilitating culture to vector meaningful connections and interactions between comprehension of the text and study of the code. Many prone reading documentary texts in parallel. The textbook readers are: *Rue des contes*; *Patati et patata*; *Mots d'école*; *Chut je lis*; *A tire d'Aile*; *Croque ligne*; *Que d'Histoires*; *Valentin le magicien*; *Max, Jules et leurs copains*; *Je lis avec Dagobert*; *Libellule*; *Un monde à lire*; *Bulle*. Two other textbook readers fit this category to all intents and purposes but will be studied separately as they count as the most frequent in the sample. They are *Ribambelle*, which scaffolds stories that the pupils have already read from start to finish, and *A l'école des albums*, which proposes a two-fold faster programme of study of grapheme-phoneme correspondences.

- analytical phonics, essentially based on teaching letter-sound relationships working up from the phoneme. The texts given to pupils are specially authored to serve progression in decoding the alphabetic system, without any attempt at storyness, even though they sometimes reach out into other types of writing. Note that while these textbooks share a similar schedule to synthetic phonics, they stand out sharply on one key point: Unlike in synthetic phonics, they scaffold memorizing whole words ("sight words") before there are decodable. The textbook readers are: *Ratus*; *Lectissimo*; *Gafi le fantôme*; *Super Gafi*; *Justine et compagnie*; *Je lis avec Mona*; *Pilotis*.

- synthetic phonics, based on explicit teaching of grapheme-phoneme correspondences (working up from graphemes) excluding any memorization of whole words and only giving pupils snatches of texts formed exclusively from graphemes that they have already been taught. The textbook readers are: *Je lis, j'écris*; *Léo et Léa*; *A coup sûr*; *Planète alphas*; *Sami et Julie*; *Taoki*.

- a quasi-whole-language approach leaving little time or space for teaching grapheme-phoneme correspondences (*Dame Coca* and *Paginaire*).

A.2.8 Evaluate class climate and pupil engagement

As outlined further up, our analytical model includes several other instructional variables that empirical research has shown to be important to achievement (Hattie, 2009) and which should thus

²⁰ To complete the study of textbook usage patterns, the field-investigators interviewed the teachers on their satisfaction with the tools they use and on their way of telling parents about how they teach reading and writing and the words they employ to describe or hallmark their "method".

be controlled for here. We know for instance that the time teachers give pupils to complete tasks is not exactly the same as the time pupils spend working on them (Crahay, 1996), but that the gap between time given and time taken gets narrower as teachers manage to recruit their pupils and capture—then hold— their attention (Hamre *et al.*, 2005). This was our rationale for inquiry into classroom climate, in other words teacher–pupils relationships and pupil–pupil relationships, and the degree of pupil engagement on programme tasks (NICHHD, 2002).

Every morning, at the first playtime, the field-investigator gauged the mean class-wide behaviour against these two indicators on a four-level scale, stating ratings as either very negative, fairly negative, fairly positive or very positive.

Seven items were used as grounds for characterizing classroom climate:

- The teacher is showing sensitive and supportive behaviour (warm, friendly, calm voice, respectful language, confident attitude, a smile...)
- The teacher is showing reassuring authority (knows how to redirect and refocus, handle conflict and clashes...)
- The teacher is betraying irritability, showing anger, shouting...
- The teacher is issuing threats, severe punishments...
- The pupils are smiling, cooperative, demonstrating interest
- Several pupils are disrespectful of the teacher (not listening, answering back, etc.)
- Several pupils are fighting, over-excited, disrupting the class...

Seven other items were used to characterize pupil engagement on classwork:

- Routines and rituals: the pupils know how to navigate the different forms of work they are given. They know what they have to do. There is no floating around.
- Immersion: during tasks, the pupils are actively immersed. They are ‘at work’ and lose very little time waiting.
- End of the task: when the pupils have finished the task given, they have a choice of other activities to engage in. They know and use these options.
- Engagement: the pupils are engaged on the tasks they are given. They look attentive, focused and interested in what they are doing.
- Teaching material: task-activity resource material is ready and to hand. It is used or handed out effectively and efficiently.
- Encouragement: the teacher fosters pupil engagement by moving around the class, observing their work or encouraging them to keep going.
- Transitions: task-to-task transitions are efficiently and effectively paced²¹.

²¹ Our team came up with these 14 items making liberal use of the American CLASS (Classroom Assessment Scoring System) instrument (Pianta *et al.*, 2008).

A.3.4 Complete the inquiry by interviewing the teachers

Several closed-ended or semi-closed-ended questionnaires served to round off our descriptive characteristics of the pupils, the teachers and the teachers' practices. These questionnaires surveyed topics such as screening for struggling pupils, special aids and services provided, work given at home, uses made of children's books, uses made of textbooks, or teachers' perceived self-efficacy. All these questionnaires connect back to our hypotheses for research, including on the influence of teachers' mental conceptions of difficulty levels in school (Piquée, 2010), on the impact of passive differentiation (Sensevy, Maurice, Clanet & Murillo, 2008), on the importance of writing as a gateway to enculturation (Lahire, 1993; Bruner, 1998), on the role of perceived self-efficacy for teachers' motivation and, consequently, the effectiveness of their action²² (Bandura, 2007), etc. The responses will be compiled, collated and analyzed to construct new level-1 and level-2 variables that will be progressively included into our analytical model. Again, there is not enough room here to exhaustively inventory all the information collected. Here are some excerpts.

A.3.4.1 Pupils struggling

To complete the observations on instructional differentiation, the field-investigator polls the teacher on the aids and services that they lean on to help cater for struggling pupils. Each week of tests, the field-investigator asks which pupils get specific in-class or out-of-class support, for how much time (in hours), and by who: a CAPA-SH-E educator, a CAPA-SH-G educator, teachers delivering APC-scheme complementary tutoring, special needs assistants, speech and language therapists²³. The field-investigator records the time investment in these aids and services as cumulative number of hours for the school year. The field-investigator also polls the teacher on the accommodations they employ in-class to help cater to struggling pupils, and how often they use these accommodations: groups streamed by level or organized by needs; flexible-grouping workshops (like at nursery school), etc.

To get a better grasp and insight into the teacher's sensitivity to their pupils, the field-investigator will ask the teacher which pupils in their class they see as "struggling", worryingly at risk of under-attaining (on reading and writing). At the end of the year, the field-investigator also asks them which pupils they intend to single out to repeat the grade-year.

A.3.4.2 Writing as culture of learning

The field-investigator polls the teacher on the practices they use to transmit writing as culture of learning so as to bridge them to the practices used for code instruction or teaching to write. The field-investigator inventories the practices observed over the two weeks running up to the interview so as to stay rooted in the facts and avoid upward declarations of intent. They typically ask the teacher whether they had used any works of children's literature over the previous two weeks, and in what format: simple storytelling, deeper study, connections to other texts, rereading, other. They also ask whether the pupils have access to a reading corner or a classroom library during independent learning time. They photograph these place-spaces and then ask the teacher whether

²² "If people do not believe they can get the outcomes they want through their own action, then they have little reason to take action or face challenges" (Carré, in Bandura, 2007, p. IV)

²³ They also identify the pupils on a PPRE [individualized education programme], coming through the MDPH [local authority gateway for people with disability], or getting special inclusiveness-scheme support for French as an additional language ('CLIN' [literacy initiation], CRI [literacy remediation], none, other...).

the class has participated in one or more cultural projects that were followed up with reading and/or writing activities. The last question asked is whether, during the school year, the class uses a library outside the school, and if so, how often.

A.3.4.3 Teacher sense of self-efficacy

The teachers were administered one last questionnaire, which was designed to capture their perceived sense of professional self-efficacy²⁴ (a version adapted from the questionnaire used in Toczek-Capelle & Ambroise, 2014). The control variable built from this sub-inquiry will be introduced into the model to characterize teachers' motivation levels and processes.

This questionnaire asks each teacher to score how capably they feel they can effectuate the 20 items below, on a scale of 0 to 100.

- Five of the items deal with instructional programme planning and scheduling: *construct a progression towards gaining fluency in reading and writing; define objectives for a learning sequence; define the target skills for pupils on a given learning sequence; determine the steps and stages needed to implement a sequence; implement a planned-and-prepared learning sequence;*
- Five of the items deal with regulation of teacher–student interaction: *communicate clearly and precisely in a language the pupils will understand; adapt your interventions to pupils' responses; foster your pupils' participation; fire your pupils' interest; capture your pupils' attention.*
- Six of the items deal with instructional assessment and differentiation: *design assessments at different timepoints of learning; analyze where pupils are going wrong to determine why they are going wrong; help pupils become aware of their progress and the effort they need to put in; design remediation activities and scaffolding to consolidate what has been taught; adapt your teaching to the diversity of your pupils; differentiate for pupils that learn at different paces;*
- Four of the items deal with authority and classroom community: *maintain the classroom-community rules; manage disruptive behaviour from a rebellious pupil; cope with conflicts and clashes between pupils; establish a reassuringly facilitative classroom environment for good learning experiences*

²⁴ Sense of self-efficacy is the strength of one's belief in one's ability to succeed in what they do. People with a higher perceived self-efficacy will set themselves higher goals and demonstrate higher engagement to achieve them (Bandura, 1986, 2007; Carré, 2004).

B. Synthesis of the main findings

What follows is a synthesis of our findings obtained through statistical testing designed to explain differences in pupils' performances at the end of first grade.

Note, however, that this study is far from finished. We have not yet processed the information collected and collated at the end of second grade [or elementary-school—year one in French-system terms]) nor have we started the qualitative analyses on all the material collected over the course of this inquiry, all of which will be published elsewhere at a later date.

The research questions tackled here led us to study the didactic strategies and instructional practices of experienced schoolteachers. The results may have been different with novice schoolteachers.

B.1 The analytical strategy

The foundation of our analytical strategy was to construct statistical units that were hierarchized, i.e. interconnected via nested relationship. Some described characteristics of the pupils (their scores of assessment test, their social background, their age, etc..) while others described characteristics of the classes in which they were learning to read and write. These classes were then characterized by a separate set of class-level indicators: pupils' mean scores, teacher's years of in-service experience, teacher's instructional practices, etc. This led us to use multilevel modelling, where the same model can mobilize the observations made at individual level (level 1) and observations made at the class level (level 2) influencing these level-1 factors.

✓ *Step one*

In the first step of the approach, we split the total variance on pupils' score at the end of first grade (global score and the sub-scores in coding, writing, and comprehension) into an intra-class variance and an inter-class variance. Analysis of inter-class variance was primary objective of this research report.

We found that whatever the test considered, the big share of variance was intra-class. That said, the inter-class share of variance was far from negligible, at an estimated 12.8% of total variance for global score here.

✓ *Step two*

We then took our multilevel models and introduced variables characterizing the pupils, essentially their socio-demographics and their initial level at the start of first grade. These variables were thought to explain much of the variance in results between pupils and between classes, i.e. to reduce intra-class share of variance and inter-class share of variance.

Reasoning in terms of pupil factors and at comparable initial levels, we found that:

- whatever the test considered, score obtained at the start of the year had very significant positive effect on end-of-year score (pupils that score higher at the start of first grade will score higher at the end of first grade);

- social background was also a driver of differences, working against pupils from low-SES backgrounds, in scores on coding, independent comprehension, listening comprehension, and writing;
- girls scored significantly better than boys on writing and slightly but significantly less well than boys on listening comprehension;
- pupils falling behind had significantly lower scores than pupils born in the first four months of the year;
- children whose home language was not French²⁵ scored significantly lower than children whose home language was French on coding, listening comprehension, and writing.

At this stage of our analysis, there was a share of inter-class variance still unexplained by the level-1 variables added to the models and assessable for effect size. Reasoning further in terms of pupil factors and initial levels, this unexplained share of inter-class variance accounted for:

- 8.06% of total variance for global score;
- 7.21% of total variance in performances on the coding test;
- 5.35% of total variance in performances on the independent comprehension test;
- 4.41% of total variance in performances on the listening comprehension test;
- 11.05% of total variance in performances on the writing test.

These inter-class variances were explainable, in part, by the teacher's practices and by the characteristics of the class in which the pupils were schooled. Our findings show that the effect size of the class factor (attending a given class) was different according to the core skills block assessed. Class had a stronger impact on writing than on comprehension (independent comprehension or listening comprehension). In other words, pupils' final performances were more strongly influenced by what happened in-class in the content area of writing than in the content area of writing comprehension.

We have formulated two complementary explanatory hypotheses for this pattern. The first hypothesis would be that performances on comprehension are more dependent on the children's linguistic and cultural baggage (and therefore family-side language socialization practices) whereas proficiencies in writing are more formal instruction-dependent. The second hypothesis would be that teachers' instructional practices are more solid, i.e. better defined, better geared and more intensive, in writing than in comprehension.

✓ **Step three**

The next step involved constructing models integrating pupil factors and their initial level as level-1 effects and the class factor (attending a given class) as a level-2 effect. These models enabled reasoning in terms of given pupil factors and initial levels and to rank the 131 classes in our population sample as a function of their net mean effectiveness. The gaps between classes were relatively small, likely because all the teachers were experienced and knew how to instil the kind of classroom climate that works for pupils' reading and writing education (see section B.8 further down). However, our calculations made it possible to single out significant classes and extract those in the sample that stood out as most and least effective for each of the skills tests administered. This step of inquiry found that it was not the same classes that were most effective in all skills areas across the board (coding, comprehension, writing), nor the least effective.

²⁵ Note, again, that information was collected from the teachers, not the families.

✓ *Step four*

Next, the last step was to explain what, in the classroom factors and the in-class practices implemented, drives these effectiveness differentials. To structure our reasoning around comparable class and teacher factors, we added six control variables to our model analyses:

- five level-2 variables factoring the classes: inside or outside the priority-education setting, density of the schoolweek (4-day week or 4.5-day week), average initial level of the class, heterogeneity of the class, social mix of the class (percentage of pupils from high-SES backgrounds);
- one level-2 variable factoring the teachers: years of service in first-grade teaching.

We thus tested our explanatory hypotheses, which had been grounded in evidence from earlier scientific research (see A.2), by taking our multilevel models and introducing the variables characterizing the instructional practices. Here, then, the phrase “all other things being equal” means after having modelled pupils’ initial level and pupils’ socio-demographic factors together with these six control variables tied to class and teacher factors.

We screened for different types of effects: a mean effect of the explanatory variable “instructional practice” on the response variables, an interaction effect between pupil’s initial level and the explanatory variable “instructional practice” and, where appropriate, an effect signposting a quadratic nonlinear (U-shaped or inverted U-shaped) relationship between the explanatory variable “instructional practice” and the response variable.

Here we report the main findings of these analyses by successively addressing the different didactic and instructional variables characterizing the teaching practices that we managed to describe and encode.

B.2 Time on instruction

Despite the strong inter-class disparities, we did not find any global effect of extending time on reading–writing instruction on pupils’ quality of learning after studying our cohort (of 2507 pupils) as a whole, nor did we find any effect after studying the times on instruction globally allocated to three out of the five big subsets of task types that we had determined: **phonological awareness**, **reading** and **writing**. However, extending time on teaching **study of language** and, to a lesser extent, **comprehension** had a positive effect on pupils’ global performances on reading and writing at the end of first grade. For the study of language, the effect is stronger for weaker pupils at the start the year (on global score).

Table 1 — 31 types of tasks proposed to pupils in first-grade classes

<p>PA: Phonological awareness</p> <p>PA1. Study the phonemes (no written product)</p> <p>PA2. Study the spoken syllables or other units of speech bigger than the phoneme</p> <p>PA3. Study the letters (name, letter formation)</p> <p>PA4. Study the correspondences between phonemes and graphemes</p> <p>PA5. Study blending or work on syllables</p> <p>R: Reading</p> <p>R1. Read silently</p> <p>R2. Recognize a whole word</p> <p>R3. Decode a word</p> <p>R4. Read aloud</p> <p>R5. Listen to the teacher read aloud</p> <p>C: Comprehension</p> <p>C1. Define or explicitly verbalize a reading intentionality</p> <p>C2. Infer, formulate or verify hypotheses</p> <p>C3. Describe and comment on an illustration</p> <p>C4. Explain or reformulate the meaning, or recount a mental model</p> <p>C5 Produce a (full or partial) recall of narrative, explanatory text or composition</p> <p>C6. Make implicit information explicit</p>	<p>C7. Propose, debate or negotiate an interpretation/various interpretations</p> <p>C8. Complete a written task involving (explicit or implicit) comprehension</p> <p>C9. Correct a written task on comprehension</p> <p>W: Writing</p> <p>W1. Handwrite</p> <p>W2. Copy (with prompt)</p> <p>W3. Copy after taking the prompt away (copy from memory)</p> <p>W4. Write from memory sentences dictated by the teacher</p> <p>W5. Produce writing by blending linguistic units already in print</p> <p>W6. Produce writing by dictating to another</p> <p>W7. Produce writing by self-encoding</p> <p>W8. Define, plan or organize the writing task (aim, who for, content, plan...)</p> <p>W9. Review their written product: comment it, correct it, revise it, improve it.</p> <p>SoL: Study of language</p> <p>SoL1. Lexis</p> <p>SoL2. Syntax</p> <p>SoL3. Morphology (word spelling, agreement declension and verbal inflectional suffixes)</p>
---	--

Drilling down to teacher-led instructional time (i.e. excluding time on work done independently or with guidance from another adult), we found it had a significant and positive effect for the initially low-achieving pupils, who showed benefits of extending time on reading–writing instruction. Reframed, this means that low achievers are more dependent on time spent with the teacher on reading–writing tasks than their peers. This influence is visible in four of the five core content areas of teaching to read and write, i.e. in phonological awareness, reading, study of language, and comprehension, but not in writing.

At this point in the study, given that we have not yet distilled the data captured in the video recordings, we can only put forward tentative hypotheses to explain why this is: time spent with the teacher holds different quality to time spent without the teacher; time spent with the teacher close fosters better interactions and better learning through the feedback received (Giasson, 1996); pupils engage more in tasks in teacher-led settings than in seatwork (Rosenshine, 1980); pupils attend at greater rates when their teacher is walking around the classroom (Berliner, 1990), etc.

These hypotheses gain strength from other results reported further down (see B.8) that argue for the positive effect, for these same initially low-achieving pupils, of settings in which teachers get strong pupil engagement in classroom tasks.

B.3 Code instruction

In the content area of command of coding and blending, we had set three objectives:

- assess pupils' performances when joining first grade, at the end of first grade, and at the end of second grade;
- identify the code instruction practices implemented in the classes;
- ascertain the effectiveness of these instructional practices.

B.3.1 Pupils' performances

✓ *At the start of first-grade*

Although the vast majority of pupils can name the letters of the alphabet, 20% can only name less than 20 of the 26 letters.

Attainment levels on the phonological skills test are modest, at around 30% on average. Pupils' performances were extremely contrasted between a minority of pupils who were already well ahead and a big half who were not at all fluent. The syllables and phonemes attack test that we used is, admittedly, a very demanding test and a potent predictor of quality of reading ability at all later levels.

The same type of distribution emerges again in the word and pseudoword reading test: at the start of first grade, 10% of pupils can practically decode whereas a good half cannot recognize any more than 2 out of a list of 35 familiar words.

✓ *At the end of first-grade*

The distribution of scores on reading a passage aloud gives a fairly good picture of the disparities in pupils' performances at the end of first grade. The average score is 39 words read correctly in the space of one minute, but a bottom 20% of pupils did not get past 15 words (averaging 8) whereas a top 20% read over 60 (averaging 81).

These results are consistent with the pattern of results obtained on the word reading and pseudoword attack test: whereas the average score is 59 points²⁶, a bottom 20% of pupils do not get past 45 (averaging 28 points) while a top 20% get over 72 points (averaging 74 points) in the time-window given, i.e. two times 60 seconds. Note that at the end of first grade, the pupils make fewer mistakes on reading frequent words (regular or exception words) than pseudowords, which points to a lexicality effect.

Performances on phonological skills confirm this trend. The average score is 22 out of 34 items successfully decoded, but a bottom [30%] of pupils did not get past 18 items (averaging 9 successes) whereas a top 30% successfully decoded over 30 (averaging 32). The bottom 30% have under-developed synthetic phonics awareness, and at the end of first grade, in one out of two cases they still do not manage to delete the first syllable of a three-syllable pseudoword.

✓ *At the end of second-grade*

²⁶ The score is computed in points and not number of words, as certain words score double when they mobilize knowledge of a positional rule (the value of the g in front of the e or in front of the a, for example).

Reading aloud again gives a good indication of the degree of command on decoding skills. The pupils' progressions are sizeable: they have gone from reading an average 39 words at the end of first grade to 76 at the end of second grade, i.e. nearly twice as many words. However, there is still huge disparity: the bottom 20% do not get past 50 words (averaging 35) whereas the top 20% read over 100 (averaging 123), which means that the fastest readers are reading almost four times faster than the slowest readers.

If we step back to remember that at the end of first grade, the bottom 20% of pupils did not get past 15 words (averaging 8), it becomes clear that the progresses made concern the entire cohort, with low achievers making huge progress. The low achievers, when not held back to repeat first grade, quadrupled their performances (jumping from 8 up to 35 words in the space of a minute) while the top achievers progressed from 81 words up to 123.

However, there is still 10% of pupils demonstrating a level of performance that is way below par, as they still score less than the all-sample average from a year earlier (less than 40 words in the space of a minute).

The word reading and pseudoword attack test confirms this result. The first decile, i.e. the lowest achievers, is clearly divorced from the rest of the population: first-decile pupils only successfully get 2 out of 3 items in the time-window given.

B.3.2 Teachers' practices

✓ *Time allocated to different kinds of tasks*

Over the course of the year, the first-grade teachers allocated an average 1h39 of teaching time per week to alphabetic system instruction²⁷, i.e. blending (81 minutes) and decoding words (18 minutes). If we add the writing tasks that involve pupils encoding speech units (54 minutes), then broad coding study (encoding + decoding) takes up over 2h30 per week. There is nevertheless huge inter-class variance, since the 25% of teachers that devote the most time do twice as much as the 25% that spend the least time.

Teachers allocate an average 38 minutes of teaching time per week to reading aloud but, here again, with big inter-class disparities reflected in a standard deviation of 18 minutes. In total, the pupils spend 3h11—i.e. 43% of their weekly learning time given to reading—writing (7h22)—on studying the basic units of speech code and written code and on putting the two together. This superset of alphabetic code system-oriented activities is far and away the leading component of instructional time on reading and writing in first grade.

✓ *Programme-of-study schedule for code instruction*

With a handful of exceptions, first-grade teachers begin explicitly teaching grapheme-phoneme correspondence right from the start of the year. At week 10 of class (i.e. by mid-November), they have already had their pupils work on an average 11.4 GPCs (SD = 4.2). Those adopting the slowest tempo (SD below the mean, i.e. 21 out of 131 teachers) had only taught 5, 6 or 7. Those adopting a fast tempo (SD above the mean, i.e. 24 teachers) had taught 16 GPCs or more.

²⁷ Weekly teaching time allocation peaks at 2h07 in the first term before dropping in the second term and dropping again in the third term.

The GPCs chosen, give or take a few, are those recommended by psycholinguists based on analysis of the French linguistic system (frequencies, regularities, etc.) (Sprenger-Charolles, 2011).

For half of the teachers, the route in is phonemic-first: they start work on segmenting phonemes before introducing the corresponding graphemes later on. For the other half, the route in is essentially graphemic-first, as they introduce the graphemes and their conventional sound values simultaneously.

✓ **Choice of text resources scaffolding shared group reading sessions** (sessions dubbed “discovering text”)

Depending on which methodological approach they adhere to, the teachers use different criteria to choose their text resources for scaffolding shared group reading sessions: some select based on value of the narrative while others select based on the grapheme-phoneme correspondences that the text contains and can be singled out for learning. We observed a broad and diverse gradient of intermediate situations and trade-offs between these two policies. If we factor in the number and nature of the correspondences taught, we find that the pupils are challenged with more or less readily decodable texts. At the start of week 10 of class, the texts provided count an average 43% of graphemes that are decodable (i.e. that they have already been taught). However, this figure can range from 11% to 76% from class to class, which means there is not at all the same individual or class-level intellectual demand on the pupils.

✓ **Textbook choices**

Our population sample counted 40 teachers (i.e. 31% of the sample) who did not use textbook readers. The remaining 91 (i.e. 69% of the sample) used thirty-odd different textbooks: 66 teachers (50%) used textbooks employing what qualified as an embedded approach, 12 (9%) used a textbook employing an analytical phonics approach, essentially based on teaching letter-sound relationships working up from the phoneme, and 13 (10%) used a textbook employing synthetic phonics, based on explicit teaching of grapheme-phoneme correspondences (working up from graphemes) excluding any memorization of whole words and only giving pupils texts formed from graphemes that they have already been taught.

We were able to verify that programme-of-study schedule for code instruction varied with type of textbook chosen. The teachers using synthetic phonics textbooks split from their colleagues by proposing a faster-than-average study of GPCs (13.8 GPCs against an average of 11.4). That said, by no means did they hold a monopoly of the fast tempo, which was also adopted by the 10 users of the embedded-approach *À l'école des albums* textbook (14.6 GPCs).

The teachers adopting a fast tempo were also the teachers providing the most readily decodable texts, although in decodability terms it was the synthetic phonics textbooks that were clearly out in front (averaging 64% decodable graphemes).

B.3.3 Influence of practices on pupils' performances

Our study emerged a number of significant effects of teachers' instructional practices on pupils' final performances.

✓ **Pace of grapheme-phoneme correspondence learning (tempo)**

A fast tempo proved beneficial to pupils' attainment in coding and writing. In coding, the influence peaked with a tempo of 14 or 15 GPCs studied in the first nine weeks of class. Initially low-achieving

pupils showed greater progression on coding when tempo was between 12 and 14 GPCs. The slowest tempos, at less than 8 GPCs, slowed pupils learning growth in coding, obviously, but also in writing.

This is a counterintuitive finding for many teachers who think the programme should not go too fast with coding instruction so as not to hold up the initially underperforming pupils. Nevertheless, our study demonstrates the counter-evidence: when initially underperforming pupils do not have enough to go on to successfully decode the written text proposed in-classroom, they make less progression. This serves as confirmation for French of the evidence in English synthesized by the US National Reading Panel (2000). Choosing a fast tempo benefits learning, as it grows pupils' cognitive clarity and self-learning ability, while at the same time cutting short room for discouragement and groping in the dark. By selecting text resources that articulate GPCs taught and proposing decoding tasks that are within their pupils' reach, the teachers facilitate their attainment and mobilize their attention and memorization while at the same time developing their sense of competency.

B.3.4 Characteristics of text resources scaffolding shared group reading sessions

There are two key results to take away, although the effects observed operate at different thresholds depending on the pupil audiences targeted.

- Teachers that propose texts with too little decodable content (less than 29% decodable) are less effective than their colleagues when looking through the lens of pupils' global scores on reading and writing.
- Teachers that ask pupils to read texts counting over 57% decodable graphemes are more effective with the pupils that are initially poor coders.

Percentage of decodable graphemes is dependent on two main parameters. It is first and foremost dependent on instructional tempo: the higher the number of grapheme-phoneme correspondences taught, the higher the probability of encountering decodable graphemes. It is also dependent on the text and the graphemes it counts. Our findings could therefore have implications for practice in the choice of text resources scaffolding shared group reading sessions, and encourage teachers and textbook authors to be more attentive to this decodability issue if they want pupils to be able to readily decode a bigger share of the texts they are given to read.

These results can only add further evidence-based legitimacy to the guidance given in the national curriculum for the basic skills stage of elementary school published in autumn 2015, which asks first-grade teachers to focus comprehension one, "like in the last year of pre-school, on texts read out by the adult educator", and two, "on texts that pupils learn to discover independently, for themselves²⁸". This second category of text (which is to be "short, easy to decode, and simple in terms of language structure and cultural referents") is recommended for leading shared group reading sessions to scaffold "the articulation between word identification processes and access to the meaning of whole sentences".

²⁸ NOR [French legislation identifier]: MENE1526483A, ministerial order dated 9/11/2015; OJ of 24/11/2015, MENESR–DGESCO. Annexe 1—National curriculum for the basic skills stage of elementary school (cycle 2) http://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=94753

✓ ***Positive influence of encoding activities***

We constructed a number of variable to capture and translate encoding tasks, i.e. moments where the pupils transcribe speech sounds into word forms, regardless of the size of linguistic units mobilized (letters, syllables, words, sentences, or texts). Two encoding task activities have a positive effect on pupils' final performances: time allocated to writing from memory sentences dictated by the teacher (where it is the teacher who chooses what to write) and time allocated to independent encoding (where it is the pupils who choose) both have a significant and positive effect on pupils' decoding performances. The effect grows up to a ceiling of 40 minutes per week for writing from memory (teacher dictates) and 35 minutes per week for independent encoding (pupils chooses). While the initially low-achieving pupils do not significantly gain from time allocated to independent encoding activities, it is they who nevertheless benefit most from extending instructional time on writing from memory sentences dictated by the teacher.

Here again, our results prop up the policy directions of the 2015 curriculum programme, which states that “reading and writing are intimately linked—well-articulated practice of these two activities consolidates the effectiveness of both”. For first grade, the programme prescribe systematic and structured grapheme-phoneme code and blending instruction “giving pupils all the time they need to practice these skills” and associating this instruction “with writing activities: encoding to get them using what they have learnt and copywork to foster orthographic memorization”. As a key takeaway from this research, we would add that encoding activities, associated with decoding activities, foster the learning of grapheme-phoneme correspondences and spelling patterns.

✓ ***Graphemes-first versus phonemes-first approaches***

We found no effect of the variable characterizing which way round grapheme-phoneme correspondence instruction leaned. Whether the teachers start work on segmenting phonemes before introducing the corresponding graphemes later on or introduce the graphemes and their conventional sound values simultaneously, the effectiveness came out the same.

✓ ***Reading aloud***

Time given to reading aloud has a significant and positive effect on coding and writing performances for all pupils, especially pupils who start out behind in the skills block assessed. In coding, the effect grows up to a ceiling of 55 minutes per week, but the difference is only significant from a floor of 30 minutes per week.

✓ ***No textbook effect***

We found no effect of textbook choices when tested in six method-based textbook categories (synthetic phonics, analytical phonics, etc.) or in three super-categories including one pooling the teachers using code-oriented textbooks (synthetic phonics and analytical phonics combined, i.e. 19% of the sample).

The ‘textbook’ variable cannot explain between-teacher differences in effectiveness, probably because experienced teachers use the same textbook in very different ways. We saw, for instance, that the tasks they propose are not the same—not in nature, nor in task time, nor in schedule, nor in the format for getting them done.

B.4 Teaching for comprehension

In the comprehension skills block, we had set the same three objectives:

- assess pupils' performances when joining first grade, at the end of first grade, and at the end of second grade;
- identify the comprehension instruction practices implemented in the classes;
- ascertain the effectiveness of these instructional practices.

B.4.1 Pupils' performances

✓ **At the start of first-grade**

As outlined further up, the comprehension skills of the pupils joining first grade were assessed via three tests: knowledge of vocabulary, comprehension of short sentences and comprehension of three texts read out. The test data gathered together points out that while the vast majority of pupils have a good knowledge of the vocabulary tested (80% of scores above the grade average) and a good understanding of short yet syntactically relatively complex sentences read out to them (90% of scores above the grade average), they suffered on listening comprehension. Analysis of the results on this test found a mean attainment score of 48.6% and emerged that the bottom 10% of low achievers fail to attain a 15% score.

✓ **At the end of first-grade**

- What do pupils understand from text they do not have to decode?

Comparison of pretest (start of the year) versus posttest (end of the year) scores on the exact same listening comprehension texts, i.e. *Anatole* and *Le bel oiseau de pluie*²⁹, shows a progression, from 44.1% attainment at year-start up to 56.6% attainment at year-end. However, a sizeable number of schoolchildren still struggle to understand these relatively short narratives even when they do not have to decode them.

- What do pupils understand from text they have to decode?

To answer this question, we analyzed the scores on two tests: sentence comprehension and text comprehension. The vast majority showed a good understanding of isolated sentences, even syntactically complex sentences, as 80% attained scores above the grade average. However, these performances tailed away for the second test (a novel test of text comprehension by independent reading), for which attainment averaged just 52.2%. Note that the top 10% achievers on this test got perfect or near-perfect scores (14 or 15 out of 15) whereas the bottom 10% got an average score of less than or equal to 1 point.

Drilling down to the two sub-tests of this text comprehension exercise (first, producing the illustration, and second, the questionnaire), we find that the pupils scored better, on average, on the first test (at 57.75%) than on the second test (48.77%) that required them to make more complex inferences.

²⁹ These two texts were singled out on the grounds that their comprehension assessment had more discriminant power than the third text and we thus better geared for charting pupils' progressions.

- What drives the quality of independent comprehension?

Regression analyses were conducted to explain the quality of independent reading that we had assessed at the end of first grade using a test (independent reading skills—see earlier) requiring deep processing of implicit content. Three explanatory variables emerged: quality of decoding, phonological skills, and listening comprehension. In other words, for the third variable, the better the pupils understand text that is read to them, the better they understand text that they read alone.

The importance of decoding for independent reading comprehension is long established, yet the importance of listening comprehension skills is rarely factored in. However, its huge factor weight (accounting for three quarters of the weight of effect for decoding) argues strongly for delivering teaching for comprehension of texts that the adult reads out loud—in nursery school, obviously, but also in first grade and second grade.

✓ **What do pupils understand at the end of second-grade?**

The two tests that we selected to measure comprehension of text read aloud (*Les enfants et la sorcière* and *Pourquoi les éléphants ont-ils une trompe?*) come from assessments proposed by the DEPP [directorate for education programme assessment, performance and policy] for a national representative sample. Our analysis of the test results showed that the scores in our sample mirror the scores obtained by the DEPP for the same tests, averaging out at just over 60% attainment, although with the same kind of caveats as for the end of first grade, i.e. that there is persistently strong inter-individual variability and that it was of course the complex questions, requiring mental inferencing, that pupils struggled with most.

B.4.2 Teachers' practices

Our analyses first focused on the net total instructional time allocated to comprehension skills and then, more specifically, on the time given to each of the nine tasks types previously inventoried (see B.2.3).

We found that the time given to comprehension was well short of the time given to coding or writing, averaging just 68 minutes per week, i.e. 15.5% of net total time on reading–writing instruction.

Inter-class disparity was huge, with a nine-fold time difference between the top and bottom deciles.

Over the course of the year, time given to comprehension was very unevenly distributed across the nine types of tasks (C1 to C9).

- Individual written tasks (C8) took up a big share of comprehension-work time, amounting to an average of 25 minutes per week.
- Oral tasks on learning to construct meaning (C3 to C7³⁰) barely took up over 30 minutes per week on average.

³⁰ Describe and comment on an illustration—Explain or reformulate the meaning, or recount a mental model—Produce a (full or partial) recall of narrative, explanatory text or composition—Make implicit information explicit—Propose, debate or negotiate an interpretation/various interpretations.

Certain oral tasks were never observed in nearly half of the classes, including tasks designed to help pupils learn to make implicit information contained in the text explicit by generating inferences (C6) and to propose, debate or negotiate an interpretation (C7), which we given little more than 8 minutes per week on average.

B.4.3 Influence of practices on pupils' performances

Despite the global under-allocation of time on comprehension, and allowing for the strong inter-teacher variability, several effects were tested.

First off, it emerged that cumulative instructional time input on comprehension tasks, taken independently, has no effect on pupils' progressions in this content area. In other words, just allocating more time to comprehension instruction is not enough to bring about significant changes in comprehension.

That said, certain temporal sequencing configurations prove more effective than others. The way teachers schedule comprehension-task sessions over the course of the year and the tasks they prioritize have an effect on pupils' learning growth. We observed that teachers who give oral tasks C1 to C7 constantly over the course the year, like teachers who give these same tasks progressively more time as the terms go by, were found to bring about positive effects on the listening comprehension test at the end of first grade in the initially low-achieving and intermediate-level pupils.

We also observed that intervention to increase the time given to these comprehension tasks over the course of the year brings about significant progress on the independent reading test in the initially poor comprehenders.

We further observed a significant effect of time spent studying language: the pupils who, starting out low achievers, get more lexis instruction show greater growth on listening comprehension, whereas it is morphology instruction that has positive effect on growth in independent comprehension. We finally observed that the number of books read in-class has a positive influence on listening comprehension (see B.6).

B.4.4 Discussion

Even if certain instructional practice configurations appear more effective than others, there are still many pupils underperforming on comprehension, both in our sample and the *DEPP* sample, against National Education Inspectorate criteria standards³¹.

This result is explainable by comparing the practices observed as reported here against the practices that the scientific literature presents as effective (*cf.* our inventory of the characteristics of these practices given under A.3.2 and the inventory put together by Maryse Bianco in the executive brief as groundwork to the consensus conference of March 2016). Not enough time spent teaching for comprehension is one factor, but the teachers' didactic strategies has to be another. Practitioner research has inventoried crucial ingredients for effective intervention, but these ingredients were short for half of the teachers observed here and altogether devoid from other half. For example,

³¹ <http://www.education.gouv.fr/cid75316/bilan-de-la-mise-en-oeuvre-des-programmes-issus-de-la-reforme-de-l-ecole-primaire-de-2008.html>

there is conclusive evidence that growth in comprehension is strongly correlated to explicit instruction and that this effect is sharper for children starting out with low comprehension scores. As the children spend more time on child-managed (not teacher-led) reading comprehension activities, the effectiveness gaps grow wider, including when these tasks are comprehension-oriented: silent reading, seatwork and decoding unfamiliar words (Connor, Morrison & Kaich, 2004). Unfortunately, this is precisely what we observed here: in practice, the pupils spend a lot of time on individual reading-comprehension tasks, and often without the teacher leading, whereas moments where pupils are encouraged to explain or reformulate meaning or recount a mental model (C4) or to produce a narrative recall (C5) only come to an average of 19 minutes per week.

We now have to understand the reasons why first-grade comprehension instruction is so weak. Is it because comprehension instruction has less spectacular and gratifying short-term effects than decoding instruction? Is it because the teachers make command of coding a prerequisite and consequently postpone comprehension instruction to later grade years? Or is it because they lack didactic resource tools capable of helping them? One thing is sure: working to an 'embedded'-approach textbook does not necessarily lead teachers to spend more time on comprehension instruction or to implement special instructional strategies.

B.5 Teaching for writing

In the writing skills block, we had set the same three objectives:

- assess pupils' performances when joining first grade, at the end of first grade, and at the end of second grade;
- identify the writing instruction practices implemented in the classes;
- ascertain the effectiveness of these instructional practices.

B.5.1 Pupils' performances

✓ Key results in first-grade

At the start of the year, a large proportion of the first graders write out three words (*rat*, *lapin*, *éléphant*) making no attempt to transcribe the corresponding speech form (46%), which means only one out of two pupils has understood the sounds-to-symbols dimension of writing. Most of the other pupils did not follow the alphabetic principle to the letter (34%), leaving only a minority (20%, mostly pupils from higher-SES backgrounds or already ahead for their age) who follow grapheme-phoneme correspondences and produce a writing that readably transcribes the blended sound sequence. In September, 15.2% of pupils failed to write the task sentence (*Tom joue avec le rat*), and 61% wrote it in a single unsegmented block.

At the end of first grade, when re-tested on these three words, the majority of pupils (67 %) employed writing that was relatively GPC code-compliant, or even readably spelling-compliant (28%). The pupils that were still struggling to transcribe the blended sound sequence (5%) were mostly pupils from low-SES backgrounds or pupils already 'falling behind'. Approaching the ability to write the first sentence (*Tom joue avec le rat*), practically all the pupils segment the sentence into words and had anchored the permanence of the written word *rat*, which is written the same way in the test sentence and in outside contexts. For the second sentence dictated to them (*Les lapins courent vite*), three quarters of the pupils did not use any plural markers at all, 8% of pupils show

premises of using them, but on the wrong kind of grammatical word class (using a plural marker for nouns on the verb, e.g. *les lapin cours vite*), 17% of pupils produced one of the two markers required (e.g. *les lapins cour vite* or *les lapin courent vites*), and only 1% of pupils got the right plural markers on both noun and verb.

Focusing on the production of text from images, most pupils produce writing in response to the instruction provided: over 85% prove capable of writing a 60-letter-minimum text in the 15-minute task window set. The readability scores show that over 70% of pupils produce writing that is fluently decodable out loud. Moving forward with the segmentation challenge, the boundarying of semantic and syntactic units, even if still clumsy at this stage, shows early premises of growing textual coherency. The large majority of pupils appear conversant with the principle of storybook narrative and already show good early knowledge of the characteristics of written narrative. A large number of pupils (over 40%) did not stop at successively describing the content of the images but seem to have understood that telling a story means going a whole lot further: over 20% added up to three extra bits of information. First-grade pupils can therefore handle more demanding pathways into narrative than they are credited for.

✓ **Key results at the end of second grade**

In second grade, in tests on selecting correct spelling, like in tests on writing dictated words and sentences, there is dramatic heterogeneity in pupils' performances, with 10%–20% of pupils posting very poor results. If we take only the markers of noun-phrase and verb-phrase number, then 30% of the sample would qualify as poor (average score of less than or equal to just 2 out of 12).

The second-grade test on text production is more complex than the first-grade test in that the images do not give any narrative storyline. The illustration only features the characters, and the task instruction—a free-form composition—is to write a story by choosing one or two of the characters. In this test again, there is equally strong heterogeneity in performances, even though a majority of pupils manifest firm appropriation of key skills for self-led text production, chiefly a capacity to handle some of the crucial elements of a written narrative (characters, places, stages in the storyline, markers of coherence and unity, etc.). For example, pupils appear to have a better command of the use of spatial and chronological sequence connectors than at the end of first grade, as they make greater and more accurate use of connectors other than *et* ['and']. On a final note, even though 80% of pupils produce an entirely readable text and therefore have a command of penmanship, they still struggle with verb agreement declension.

B.5.2 Teachers' practices

Writing tasks are given every week for 2h23, i.e. practically a third of the 7h22 of teaching time given to reading–writing. Observational analysis showed that copying gets the biggest share of time spent on writing (34% of time on writing, 47 minutes per week), followed by producing writing (29% of time on writing), where the 42 minutes per week budget is essentially spent on producing written text independently (25 minutes) and producing written from pre-printed linguistic units (12 minutes).

On average, the time given to writing sentences dictated by the teacher comes to 16% of time given to writing (22 minutes). Handwriting (11% of time on writing, 16 minutes) and planning and revising texts (10% of time on writing, 14 minutes) are where pupils spend the least time.

Looking at the types of linguistic unit worked on in writing tasks, the results suggest that from the teachers' perspective, there are linguistic units that fit certain tasks well (the letter works for

handwriting tasks, the syllable works for writing dictated sentences) whereas other units can be tackled via different tasks (sentence and text for both copying and composition). From this standpoint, the word pans out as the linguistic unit worked on in the biggest number of different tasks (copying, writing out dictated sentences, and composition). Text-as-unit gets worked on in both copying and composition, more so in the second half of the year, but is still a minority practice: 7 out of 131 classes in November, 38 in March and 45 in May got to try their hand at independently composing and producing written text. This finding shows that the work is initiated in first grade even if the actual time allocated to it remains next to none (less than 2 minutes in November; around 8 minutes in March; 10 minutes in May).

B.5.3 Influence of certain writing tasks on pupils' performances

Cumulative time on all-aggregate writing tasks has no effect on pupils' final performances. However, certain writing tasks do improve pupils' performances.

✓ *Write from memory sentences dictated by the teacher (W4)*

Increasing time on write-from-memory tasks has a positive effect on the writing performances of initially poor writers, and this positive effect is also visible in their scores on writing from memory and on producing text.

Increasing the time spent writing from memory has a positive effect on the writing performances of intermediate-level writers, and particularly on end-of-year write-from-memory scores. However, there is a ceiling effect: spending more than 40 minutes per week on this task activity does not bring any further progression.

In summary, spending time writing dictated sentences grows writing development primarily for initially low-achieving and intermediate-level pupils, i.e. 42% of the pupils on our sample, with a ceiling effect capped at 39 minutes per week.

✓ *Writing production (W5 to W9)*

Cumulative time on tasks involving producing writing, including planning and revising tasks (i.e. the aggregate of W5 to W9), has a positive effect on the write-from-memory scores of the bottom 11.5% pupils furthest behind.

If we remodel the effects of each of these tasks separately, we find that spending time producing writing by blending linguistic units already in print (W5) has a negative global effect on the final writing performances of the pupils at all levels (from the lowest to intermediate and highest achievers). This task ultimately proves counterproductive regardless of the pupils' initial levels of writing attainment.

That said, spending time teaching pupils to plan the writing task and review their written product (W8 and W9) has a significant positive effect on comprehension scores at the end of first grade. This also has an effect on writing scores, especially for the weakest writers.

The writing production tasks that involve encoding pupil-chosen linguistic units (W7) work well for all pupils, especially the 805 initially strong pupils (32.4% of the sample).

✓ *Number of explicit cues*

Number of tasks prompting the teacher to give at least one explicit cue during writing-task activities on shorter linguistic units (syllables and words) has a positive effect on the writing

development growth of the weakest writers. For tasks involving producing sentences and texts (W5-W6-W7) and tasks involving planning and revising texts (W8 and W9), this same number has effects on the strongest pupils' performances. It also has a significant effect on comprehension scores.

Certain writing tasks, including child-managed production of writing and (teacher-led) writing from memory, appear to carry more weight than others, with contrasted effects depending on the pupils' initial skills levels.

B.5.4 Catalysts for further inquiry

✓ *Marking noun-phrase and verb-phrase plurals*

The pupils' test results show that the acquisition of plural-marking in a spelling task (writing dictated sentences) is not yet attained at the end of first grade. It is only accessible to a tiny fraction of the sample. At the end of second grade, just 30% of pupils command the markers of noun-phrase and verb-phrase number. However, time allocated to morphology and syntax has a positive effect on pupils' performances (see B.6).

There are no fast-tracks or short-cuts to the acquisition of plural-marking and agreement, so it makes sense start out early, probably right from first grade.

We would advocate a deep didactic rethink of the basic-skills-stage curriculum (cycle 2) to better accommodate the time needed to learn the markers of noun-phrase and verb-phrase number.

✓ *Writing–spelling linkage*

Agreement declension may be difficult in spelling tasks (writing dictated sentences), but it is even more so in writing production. However, our results show that yes, agreement takes time to learn, but they also show that first-grade pupils are already capable of writing texts that look like texts. Consequently, they should not be made to wait until they command agreement before getting opportunities to write texts. A rethink should therefore start from how spelling is handled in writing production tasks, including on revising text, but also—crucially—from the necessity for children to produce writing, rather than just copying it, even before they have mastered coding skills.

B.6 Study of language

B.6.1 Pupils' performances

✓ *Key results in first-grade*

As reported under section B.5.1, when asked to make a noun-phrase and a verb-phrase agree, three quarters of pupils did not use any plural markers at all and the remaining quarter was split between wrong markers and only partially right markers. As we said, only 1% of pupils got the right plural markers on both noun and verb. It is therefore reasonable to conclude that, regardless of the teachers' practices, the majority of first-grade pupils were not ready to access the grammar of agreement declension in a dictated sentence task situation.

✓ *Key results in second grade*

At the end of second grade, we observed how pupils handled verb agreement declension in producing writing and in writing dictated sentences. In writing dictated sentences, 5.9% of pupils inflected the *-nt* affix marking agreement in both the sentences dictated. In production of writing, there was not one correct mark to inflect person on the last verb of the text, but 92% of pupils came up with another mark to that expected. These results are correlatable with the identification of grammatical word class: in a text counting four verbs, 12% of pupils identified all four and 30% identified three. In other words, verb identification and the recognition of verb-form markers has barely begun at the end of second grade.

Noun phrase agreement declension is observable in a spelling exercise where, depending on syntactic construction, the morphological features of the nouns and adjectives are to be inflected to make the sentence agree (e.g. [third-person subject-plus-verb] *Il descend les ...* + 4 spelling pattern choices for [the feminine plural noun phrase] *pentes enneigées*), the results were bracketed between 1/3 and 2/3 right. In writing dictated sentences, 21% of pupils inflected the plural mark on the three nouns and 8% on the three adjectives. However, if we count partial marker inflection, we again reach 2/3. The identification of three underlined nouns in a short text was complete for 23% of pupils and partial for 16%. The identification of three underlined articles was complete for 20% of pupils and partial for 7%. At the end of second grade, then, only a minority of pupils catch the noun-phrase constituency structure, with strong variations according to lexical vocabularies and syntactic constructions.

B.6.2 Teaching practices

The study of language is alive in first grade, where we counted three types of content area-relevant tasks: tasks connected to lexis (SoL1: words, what they mean and how they are formed), syntax (SoL2: sequencing words in a sentence and sentences in a text) and morphology (SoL3: spelling, agreement declension and verbal inflectional suffixes).

Global time allocated to study of language increases over the course of the year, climbing from an average 30 minutes per week in November to 40 minutes in March and 43 minutes in May. However, this time-budget follows different time-course curves depending on the linguistic unit of focus: time spent on studying lexis increases fairly constantly over the course of the year, whereas time spent on morphology-oriented tasks increases significantly from an average 4 up to 13 minutes per class, with a sharp quantitative jump in March. Study of syntax follows the reverse pattern: it starts out the biggest content area of focus at the start of first grade, chiefly as teachers introduce the notion of sentence, then gently declines down to the same level as lexis and morphology.

Our models found huge class-level disparity in weekly teaching time allocated to study of language. The first decile (set of the 13 classes that spend the least time on SoL) spend an average 7 minutes per week on the study of language whereas the tenth decile (set of the 13 classes that spend the most time on SoL) get 83 minutes.

Study-of-language tasks are more often than not integrated into sessions outside grammar, vocabulary or spelling (for instance, reading in 'discovering text' sessions, for instance). Over two thirds of SoL tasks do not add up to more than a cumulative 10 minutes and account for less than 40% of the tasks worked on during the session.

Every class features a series of metalinguistic terms in its classroom-wall materials and in the pupils' classbooks, and these terms get increasingly employed as the year goes on.

B.3.3 Effect of practices on pupils' performances

Time spent studying language has a significant and positive effect on pupils' global performances on reading and writing at the end of first grade. This result adds to a growing literature base providing evidence that early-age influence of morphological and orthographic awareness on the development of reading literacy (see, for example, Apel *et al.*, 2012; Deacon, 2012; Sanchez *et al.*, 2012).

Specifically, if we refine the analysis by stratifying per skills area, we get to the findings reported below.

- Time spent studying language has a significant and positive mean effect on the progressions of first-grade pupils in every literacy content area across the board except progression on listening comprehension. Note, however, that for writing score and global score, this effect differs in size with pupils' initial levels.
- Time spent studying lexis has a significant and positive mean effect on all end-of-year scores, except for the score independent comprehension. For listening comprehension and for coding, this effect concerns low-achieving and intermediate-level pupils.
- Time on studying syntax only has an influence on pupils' progression in writing, whereas time on studying morphology only has an influence on pupils' progression in writing and in independent comprehension.
- Time on tasks that articulate writing with morphology has a positive effect on pupils' progressions in every literacy content area considered, except listening comprehension.

On a final note, preliminary analysis of the sessions videoed in the most effective classrooms (*cf.* E.3) reveals that the teachers are particularly attentive to metalinguistic discourse: they explicitly signal the transition from using language to studying language ('language about language'), anticipate the kind of misunderstandings that lead certain pupils to interpret a word of reference referentially³², and diligently use a compact set of metaterms (*sentence, noun, verb, masculine or feminine, singular or plural*). These teachers effectuate a more reflexive transmission of skills than the least effective classes by scaffolding pupil–pupil interactions as much as teacher–pupil interactions. This prompts us to stress appropriate pre-service and ongoing in-service teacher education as critical to study-of-language practice.

B.7 Literacy enculturation

✓ **Contrasted practices**

Our early analyses made it possible to rank the classes into categories on a scale of enculturation through literacy defined on the basis of the "writings", "uses of writing and writing spaces" and "storybooks" variables. Literacy enculturation is poor in 34% of our in-sample classes, average in 37% and strong in 29%.

³² An illustrative example would be when the teacher explains that in the word *loup* you can hear the sound /u/ but a pupil answers saying "but I'm afraid of the wolf". In this situation, the pupil is talking about the thing, not about the word, i.e. referentially interpreting the noun of reference.

The classes qualified as highly enculturational are characterized by a combination of prominent use of text, efforts to provision varied types of writing, prominent and diversified use of storybook readers, and a cultural offer designed to foster self-appropriation by the pupil.

These practices are less dense in the averagely enculturational classes.

Lastly, poorly enculturational classes make little use of text and little effort to ensure a variety of types of writing. The storybook reader gets relegated down and the narrow cultural offer is not designed to foster self-appropriation by the pupil.

This categorization scheme, despite collapsing a variety of practices into each class–category, nevertheless emerges different degrees of literacy enculturation from class to class.

✓ ***Significant effects of the progression of low achievers***

When the composite variable “Literacy enculturation” is modelled through the analytical strategy set out under B.1, it emerges that the highly enculturational classes foster the progression of initially poor and intermediate-level pupils across-the-board in every area of attainment of reading–writing skills. It is this for these pupils that practices geared to literacy enculturation work best.

Literacy enculturation fuels the emergence of written competency in every one of its dimensions. It can contribute to narrowing educational inequality since it has an effect on the pupils most at risk of school failure.

Literacy enculturation, as defined here, joins up to social uses (and not just in-school uses) of reading and writing. As a discipline of study, then, literacy enculturation lies at the interface between sociology approaches and didactic approaches to reading and writing development. The positive effects identified prompt us to formulate a new hypothesis: reading–writing instruction is more effective when it articulates the enculturation of social literacy practice and didactic literacy practice. If this new hypothesis is verified, then this articulation would warrant being instructionally re-engineered on the primary teacher training side and implemented more explicitly and engagedly on the classroom side.

✓ ***Quantitative results that warrant finer-grained analysis***

We plan to press ahead on this study by mobilizing qualitative approaches, chiefly by exploiting the video evidence. The expectation is that this orientation will enable us to fully capitalize and consubstantiate this first quantitative insight: fine-grained analysis of lesson materials, variety of ways of taking from the picturebook and interventional interactions around books, independent reading time given to pupils, how classes invest reading corners, treating the early reader as “cultural object”, etc.

Qualitative follow-through is necessary as the enculturational process we aim to identify and qualify is so complex and polymorphous that it cannot be grasped in its entirety by an exclusively quantitative approach, which thus needs to be rounded off by a more detailed descriptive analysis of in-class practices.

We will thus exploit the video, photo and other traces of evidence collected to conduct a finer-grained analysis of the practices described and reported here, and thus see the focus of our observational study from a different angle.

B.8 Non-content-area-specific variables

B.8.1 Class climate and pupil engagement

After slimming down our readapted version of the American CLASS (Classroom Assessment Scoring System) instrument, we assessed classroom **climate** via a compact set of items characterizing the behaviours manifested by both teacher (“sensitive and supportive”, “smiling”, “irritable”, “angry”, etc.) and pupils (“smiling”, “disrespectful”, etc.). Despite relatively little variability, attesting to the fact that the teachers who accepted to have researchers in their classroom had all created a positive classroom climate, the results do reflect that positive classroom climate has a positive effect on pupils’ performances in reading and writing. Better class climate translates into better learning.

Even though these effects are still chronically under-studied in French schools, our results stack up with evidence from correlational studies conducted in English-language primary schools (for an integrative review, see Thapa, Cohen, Guffey & Higgins-D’Alessandro, 2013). They also bring further insight into why low-achieving pupils are the only ones for which the classroom climate effect is significant when the cohort is stratified into three subsets according to pupils’ scores at the start of first grade.

Using a set of seven items, we described the mean **level of pupil engagement** in the classroom, i.e. the pupils’ involvement in learning tasks and participation in learning activities, and the teacher-led actions for scaffolding engagement. Here again, it emerges that the initially low-achieving pupils progress more in the classes in which we observed a better mean level of pupil engagement. In other words, certain teachers manage to keep their pupils engaged on-task better than others, and the positive effect of this better engagement is stronger for weaker pupils.

In short, our results enhance current knowledge on struggling pupils by underlining how the instructional dimensions of classroom climate and engagement are every bit as important as the didactic strategies of the teachers.

B.8.2 Instructional differentiation

✓ *Effect on low achievers of the teacher’s aide and intervention*

As was the case for the variables capturing time-budget, classroom climate and pupil engagement, the variables characterizing the differentiated instruction strategies implemented by the teachers (different tasks set, extra scaffolding, etc.) do not have global effect on the full cohort of pupils. Here again, though, effects did emerge for initially low-achieving pupils.

We observed that around a quarter of reading–writing tasks were differentiated in some way. Among the various formats of instructional differentiation inventoried, the most common (used for 20% of tasks given) is to give the same task to all the pupils and provide special support to those that need it. This type of differentiation, based on differentiated scaffolding for a common activity, has positive effects on the performances of the low achievers but zero effect on the performances of the rest of the class.

However, giving certain pupils tasks with a different content to that assigned to the rest of class-group has negative effects on the results of the low achievers. This type of task proves rare, accounting, on average, for just no more than 4% of the total tasks inventoried.

✓ *Out-of-class aide and intervention*

We inventoried the various out-of-class intervention frameworks on offer.

- 26% of pupils get aid from their teacher, provisioned through the *APC*³³ (complementary tutoring) framework.
- 12.4% of pupils get aid delivered by a speech and language therapist.
- 14,9% of pupils get special aid from a member of the RASED (special education network for struggling pupils³⁴): 12.7% by the CAPA-SH-E educator, 4.5% by the CAPA-SH-G educator, 2.3 % by both.

The intervention stacks up on top of classroom work with the first two types of aid (*APC* scheme and speech and language therapist), but not the third intervention which, with a handful of exceptions, takes the child out of the regular classroom and into a special education setting. Basically, a large number of pupils get some kind of supplemental in-school intervention, be it the *APC* or the RASED, but very few get intensive intervention, i.e. averaging at least an hour per week.

The effect of these interventions is unknown, as the assessment of their effectiveness runs up against methodology problems that are tricky to resolve. For example, amongst the subset of pupils given out-of-class aides, over half co-attend more than one intervention programme, making it a tricky to disentangle their effects. Either way, we can register the fact that the supplemental interventions on offer fall well short of the international benchmarks with evidence-based effectiveness, which budget up to 40 hours and prove more useful if they are frequently-dosed (several sessions a week) and packed into a tight window of no more than 20 weeks (Elbaum *et al.*, 2000). A majority of US recovery programmes, for instance, have students in intervention for between 20 and 50 minutes per day (Wanzek & Vaughn, 2008). In short, the interventions analyzed here in France do not share the characteristics of the intensive interventions that normally have a positive effect on pupils' progressions.

On a final note, the figures also highlight the scale of use of speech and language therapists, thus confirming several previous studies reporting that a share of the reading recovery effort is being divested over to the paramedical sector (Morel, 2014; Viriot-Goeldel, 2007).

B.8.3 Explicitness of the teaching

We studied explicit instruction, in its myriad forms (spanning clarifying the goals of the task or giving cues and clues on how to tackle it), and didactic memory strategies (refreshers on earlier knowledge taught, moments for institutionalization).

✓ *Teaching practices*

Our descriptive analyses of direct in-classroom observations came out with two main patterns: one, a huge diversity of practice in explicit cueing, and two, under-use of didactic memory.

Specifically, on explicit instruction, our observations found:

- instances of explicit cues that differ wildly between classes and according to period of the year;

³³ <http://eduscol.education.fr/cid74795/les-activites-pedagogiques-complementaires.html>

³⁴ <http://www.education.gouv.fr/cid24444/les-reseaux-d-aides-specialisees-aux-eleves-en-difficulte-rased.html>

- a decrease, over the course of the year, in number of tasks prompting at least one explicit cue;
- greater frequency of tasks prompting at least one explicit cue on the “how” (i.e. on procedures and strategies) compared to tasks prompting at least one explicit cue on the “why” (goals, outcomes and challenges of the task);
- an unequal distribution in instances of explicit cues between the big five supersets of tasks, with writing tasks counting the highest number.
- Focusing on didactic memory, the analysis shows:
- a higher number of instances of tasks mobilizing of the class’s type-1 didactic memory (DM1: connecting the pupils to previous learning-work situations to guide the mobilization of knowledge they have already been taught) than type-2 didactic memory (DM2: wrapping up the session with moments to co-articulate, with the pupils, their growth in learning);
- variable use of didactic memory according to the task it articulates with;
- the importance of ‘refresher’ tasks (DM1) in the study of language, even though this task category is allocated less time on average than any other.

✓ ***Effect of practices***

We first characterized the classes as a function of number of tasks prompting at least one explicit cue, without factoring in the nature of the tasks numbered. We found that the variables characterizing number of tasks prompting at least one explicit cue on the “how” or the “why” have no mean effect on pupils’ performances, and nor do the variables characterizing mobilization of didactic memory.

The next step now is to push the data modelling forward by systematically studying the effect of explicit cues as a function of tasks and as a function of the linguistic units targeted by these explicit cues. The first data modelling led in this direction, which was to study the impact of explicit cues in writing tasks, highlights an effect of number of tasks prompting the teacher to give at least one explicit cue during writing-task activities on pupils’ progressions in coding and in comprehension. Furthermore, the “Teaching for writing” section of this synthesis (see B.5) shows that number of writing-task activities on shorter linguistic units during which the teacher gives at least one explicit articulation has some effect on the writing growth of the weakest writers. These are findings that prompt us to formulate the following hypothesis, compatible with mounting evidence from earlier scholarship (including Bautier & Goigoux, 2004; Bernstein, 1975/2007; Fijalkow, 2014), i.e. that explicit instruction has an effect on growth in reading and writing learning if and when it targets certain tasks, certain linguistic units, and certain pupils. This is a hypothesis that we will be testing in the next wave of analysis.³⁵

³⁵ The bibliography to this executive brief can be read with the full report via the IFé website, at <http://ife.ens-lyon.fr/ife/recherche/lire-ecire/rapport/rapport>

LiRe
et
ÉCRIre