Teaching Spelling with Twitter: The Effectiveness of a Collaborative Method for Teaching French Spelling

Caroline Viriot-Goeldel*° & Thierry Pagnier*°

* Univ Paris Est Creteil, Bonneuil | France

° UPL, Univ Paris 8, Saint-Denis | France

Abstract: Twictée, a portmanteau of Twitter and *dictée* (French for dictation), is a collaborative method for teaching spelling that promotes the metacognitive reasoning needed to understand and assimilate the morphosyntactic features of French spelling. The present study evaluated Twictée's impact on spelling performance in 40 classes of 4th-, 5th-, and 6th-grade students (N = 893 students). Mixed-model analyses showed a significant improvement in global spelling performance over time, but the impacts of the interaction between time and condition reached significance for only four specific aspects of spelling performance. Nevertheless, further analyses showed that Twictée's overall impact on spelling performance was significantly greater in schools in disadvantaged urban areas and in large classes. We discuss these results in the light of previous qualitative analyses carried out on this corpus.

Keywords: writing, orthography, spelling, collaborative writing, teaching

1. Context

The processes that beginning writers rely on depend on the specific features of each language, as well as on general principles common to all languages (Sprenger-Charolles, 2004). French spelling is particularly complex, mostly because of its highly semiographic nature (Jaffré, 2004). Although this complexity can be an advantage when it comes to reading, it makes spelling difficult to master. French spelling is alphabetic (graphemes, which are units of writing, represent phonemes), but correspondences between graphemes and phonemes are often opaque (Peereman, 1999; Ziegler et al., 1996), so even writers who have internalized the system of correspondences can rarely predict the correct spelling of words and must therefore memorize their forms. Spelling mistakes also result from the fact that many grammatical markers of agreement and number present in written French are not pronounced (Brissaud et al., 2014). In fact, inflectional morphology is the most frequent cause of French students' spelling mistakes, and even well-educated adults make this type of mistake (Bosse et al., 2020; Morin et al., 2018). Consequently, constant thought and attention must be paid to spelling when writing in French.

Children's ability to master these difficulties of writing in French appears to be declining, with several studies reporting a deterioration in mean spelling performance among French students (DEPP, 1996, 2016; Manesse & Cogis, 2007). Although this slide in performance has made it urgent to determine the effectiveness of tools used to teach spelling, studies (whether in psycholinguistics or didactics) examining the impact of methods for teaching French spelling (whether traditional or novel) are rare and have mostly involved small numbers of classes. International meta-analyses of orthography (Galuschka et al., 2020; Graham & Santangelo, 2014; Wiliams et al., 2017) have found few French studies on this issue, and most of those they found focused on invented spelling by young children and therefore did not cover inflectional morphology.

In 2013 a group of French primary school teachers created a new, collaborative method aimed at remediating this decline (Hobart & Forgione, 2015). Dubbed Twictée, a portmanteau of Twitter and *dictée* (French for "dictation"), this method has since aroused great interest among teachers looking for innovative and effective ways of updating the way they teach spelling. For example, more than 1,000 teachers in French-speaking primary schools and junior high schools around the world registered for the 2017-2018 Twictée season.

2. Theoretical Grounding of Twictée

Twictée is a multi-component spelling intervention that contributes both to the development of teaching practices and to an ongoing research project. Its components include writing under dictation, writing collaboratively, editing,

activities involving metacognitive discussions aimed at developing spelling consciousness, and using feedback to self-correct.

2.1 Writing Under Dictation

Spelling interventions in the English-speaking world most commonly have four dimensions: 1) work on phoneme/grapheme relationships, 2) orthographic instruction (explicit instruction and application of graphotactic or orthographic-phonological spelling rules), 3) memorization exercises (e.g., copy, cover, and compare sheets), and 4) morphological interventions (Galuschka et al., 2020; Wiliams et al., 2017). All of these dimensions focus on spelling individual words. In contrast, the above-mentioned features of the French orthographic system mean that French spelling has to be taught and practiced at the sentence level. This is traditionally done through dictation, a highly cherished practice in French schools (Chervel, 2006). Dictation involves the teacher reading a text and the students writing what they hear. Students in early grades generally start with single-word dictation, a method that has proved its efficacy (Riou, 2017), and gradually progress to sentences and longer texts as they get older. The efficacy of sentence and text dictation remains under-researched, but the way dictations are corrected (subtracting points for each mistake) has been criticized as demotivating (Jaffré, 1992). Dictation has also been disparaged for being more of an assessment tool than a learning opportunity.

These criticisms have spurred the development of numerous variations on the traditional dictation method aimed at making dictation more of an active learning exercise (Brissaud & Cogis, 2011). Thus, methods such as "zero-mistake dictation" and "today's sentence dictation", which their developers present as metacognitive and interactive, use dictation to get students to discuss spelling difficulties by explaining their spelling choices out loud, with guidance from the teacher. Some of these methods appear promising (Fisher & Nadeau, 2014; Cogis et al., 2015). Twictée, whose components include collaborative writing and editing, is another variation on the classic dictation method.

2.2 Collaborative Writing

Meta-analyses have shown that peer assistance has large overall effects on writing performance. Having peers help each other plan, draft, revise, and edit texts is an effective instruction strategy that enhances writing quality (Graham & Perin, 2007; Graham et al., 2012). However, few studies examining the impact of collaborative writing on writing performance have focused on the spelling component of collaborative writing.

Spelling discussions are a form of collaborative writing that is widely used in French- and German-speaking primary and middle schools (Geist et al., 2019; Viriot-Goeldel & Brissaud, 2019). Discussing spelling problems enables students to

collectively determine a solution by explaining to each other their knowledge of spelling rules and patterns, their strategies, and their procedures, and by applying grammatical reasoning. Results of the few studies to have tested the method's efficacy are encouraging. For example, Fisher and Nadeau's (2014) tests of two spelling-discussion interventions with 26 classes showed that students made significant progress in spelling, in both dictation and free writing, and suggest that using metalanguage in these discussions can enhance students' spelling performance. Nevertheless, further studies are needed to assess the impact of spelling discussions and the factors that moderate their efficacy. The spelling discussions included in Twictée give the method a collaborative writing component.

2.3 Developing Spelling Consciousness

Students need to have spelling rules explained to them and to practice applying these rules until they have mastered them. Applying spelling rules requires a large array of metacognitive skills that go far beyond declarative knowledge and includes "thinking about and reflecting on how to spell words and sentences, knowing which strategy can be used in a particular situation and subsequently applying these strategies correctly" (Cordewener et al., 2018, p.136). The ability to conduct metacognitive reflection on one's spelling is called "spelling consciousness" (Block & Peskowitz, 1990).

Students who are conscious (vs. not conscious) of their spelling process and who are able to evaluate the correctness of their spelling obtain higher spelling performance scores (Block & Peskowitz, 1990) and are better at choosing the most appropriate spelling strategies (Kreiner & Green, 2000). Although spelling consciousness and spelling performance are distinct skills, they are intimately related (Cordewener et al., 2018). Therefore, offering spellers opportunities to think about the spelling of each part of words and sentences may stimulate them to think actively about their spelling, improve their spelling consciousness, and thereby help them detect and correct spelling errors, even without explicit metacognitive instruction (Cordewener et al., 2018). Spelling discussions may encourage students to do this, as may activities such as editing texts. In fact, "the same skills that enable one to spell without errors are the skills necessary to recognize an error, and these are the same skills that determine if an error has been made" (Vanderswalmen et al., 2010, p.374). Being able to recognize errors is a component of spelling consciousness (Block & Peskowitz, 1990).

Trying to explain their spelling mistakes to a distant interlocutor may prompt students to think about spelling in different ways, as this process requires them to identify the orthographic rules that apply to a situation and to express these rules clearly and concisely. Tagging mistakes also requires students to think about spelling, determine the way in which a word was misspelled (e.g., incorrect subject-

verb agreement, verb incorrectly conjugated), and recall the corresponding rule. Moreover, connecting a specific mistake to a more generic orthographic rule helps students categorize orthographic phenomena.

To identify, explain, and tag spelling mistakes students must consciously manipulate orthographic terms and concepts, which may help them enhance both their orthographic knowledge and skills and their metacognitive monitoring (Vanderswalmen et al., 2010). Indeed, understanding the principles of orthography is beneficial for learners and is a key component of effective spelling interventions (Galuschka et al., 2020).

For this reason, Twictée includes four components that implicitly stimulate metacognition by providing students with a structured way of thinking about spelling. These four components are 1) participating in spelling discussions; 2) collaboratively editing other students' sentences; 3) recognizing mistakes; and 4) explaining in writing what the correct spellings should be and using a tagging process to categorize mistakes.

2.4 Using Feedback for Correction

Spelling teaching is more effective when exercises include explicit feedback (Totereau et al., 1997). Interventions that include providing feedback on spelling mistakes and using that feedback to self-correct are likely to improve spelling performance. Twictée requires students to correct their writing by reading and applying feedback from students in other classes. If writing under dictation, collaborative writing, editing (which requires identifying mistakes, explaining them to others, and tagging orthographic phenomena), and receiving feedback improve spelling skills, then spelling interventions based on these principles should improve students' spelling performance.

The current study's objective was to determine the impact of the Twictée intervention on 4th- to 6th-grade students' spelling abilities. We also compared Twictée's impact on students from advantaged backgrounds with its impact on students at schools in disadvantaged urban areas that benefit from compensatory education programs (known in France as "Priority Education" schools).

3. Method

3.1 Description of the Intervention

An intervention's design principles can be described by presenting its instructional and learning activities (Bouwer & De Smedt, 2018). In the case of Twictée, teachers collaborate remotely to produce a dictation sentence that covers the aspects of spelling they want to focus on and plan the collaboration between classes. Implementing Twictée then involves five or six steps, as described below, the final step being optional.

Table 1. Description of the Intervention

Step 1	Dictation	Teachers collaboratively create a sentence and dictate it to their students, who write the dictated text, working individually.
Step 2	Collaborative writing	Students work together in small groups to produce a version of this sentence they all agree on. To do this, they follow the principles of spelling discussions (Arabyan, 1990; Geist et al., 2020), processing one word after another, comparing their spellings, and explaining their choices using metalinguistic reasoning. The groups' final sentences are then tweeted to another class.
Step 3	Correction	Teachers provide the correct version of the dictated sentence. They show the different spellings used by students and ask them to explain their choices. This process is similar to the previous step, except for the fact that the teacher scaffolds students' reasoning and validates correct spellings.
Step 4	Collaborative editing: identifying, explaining, categorizing	Students work in groups to edit another class's group dictation. This process includes finding the mistakes in each Twictée and then providing feedback via 280-character tweets (known as "Twictools"), in which students explain the correct spelling. They also categorize mistakes using hashtags, with each type of mistake having its own hashtag. For example, #AccordGN is used when agreement within a noun-group is not respected. Students tweet their Twictools to the other class.
Step 5	Collaborative correction	Students receive another class's feedback (sent via Twictools), which they use collaboratively to correct their group dictation.
Step 6	Assessment dictation	Teachers assess students' learning by providing a final dictation of either the initial text or an appropriate new text with similar spelling features.

Figure 1. Example of a Twictool

#AccordPP (past participle does not agree with the subject of the verb).
Example: mes amies sont venu**es** ce matin (my friends came this morning)
#Twictool VENUES is written with -ES because a past participle used with the auxiliary être SONT (to be ARE) must agree with the subject MES AMIES. #Accord PP

Source: https://www.twictee.org/wp-content/uploads/2017/10/Dicobalises-V2-1.png

Twictée's main aim is to prompt students to think explicitly about how the language system works and thereby enable them to develop orthographic representations. The intervention involves collaboration on several levels, including among students—during the spelling discussion and Twictool production steps—and between classes—during the dictation exchange and correction steps and during the exchange of Twictools.[1] It also incorporates several digital tools (Twictée Association website, private tweets between teachers, Google Docs, Slack messaging site), which teachers use to collaboratively prepare and organize each Twictée. Students exchange dictations and Twictools via Twitter, but the extent to which they use these tools varies according to how the Twictée is carried out. Most students write their dictations on paper, and some students even prepare their Tweets using pen and paper. Finally, the Twitter format means that explanations must be brief and results in Twictools tending to follow more-or-less the same structure. In addition, Twictée has adopted an emblematic feature of Twitter—hashtags—to help students learn to categorize mistakes. Finally, using "techno-language elements" (Paveau, 2016: 25) such as @ symbols to introduce the names of the classes to which a Twictool is sent and hashtags to introduce the categories of mistakes gives students the opportunity to work on their computer literacy skills.

3.2 Implementation Checks

We used measures of two variables—fidelity and intensity—to determine potential differences in the way teachers implemented Twictée (Resnicow et al., 1998). The implementation fidelity measure involved determining the proportion of teachers/classes who performed each of the Twictée steps. The reliability of measured effects is greater when interventions are implemented faithfully, a condition that is met, according to Durlak and DuPre (2008), when at least 60% of an intervention's activities are performed. We measured fidelity by coding each task the teacher gave to students during a full Twictée cycle and then used these coded tasks to determine which of the six Twictée steps each class performed. Fidelity rates were the percentage of the total number of classes that performed each step. Fidelity rates were high for all of Twictée's steps, except for the assessment step, which Twictée's creators presented as optional (Table 2).

Table 2. Implementation Fidelity

Twictée Step	Fidelity rate
Step 1: Dictation	100%
Step 2: Collaborative writing	100%
Step 3: Correction	89.5%
Step 4: Collaborative editing: identifying, explaining, and categorizing	94.7%
Step 5: Collaborative correction	73.7%
Step 6: Assessment	52.6%

The measure of implementation intensity took into account the amount of time teachers allocated to the Twictée cycle we observed and the number of Twictée cycles they conducted during the academic year. In contrast to the consistently high fidelity rates, implementation intensities varied substantially. The duration of Twictée cycles differed significantly ($M = 196.95$; $SD = 73.61$; min = 68; max = 297), so we investigated these differences further in our qualitative analysis (Crinon & Viriot-Goeldel, 2021). In addition, the number of Twictée cycles performed by each class ranged from 3 to 6 ($M = 5$; $SD = 1$). However, we did not find a significant link between either of these implementation intensity indicators and overall changes in students' spelling performance ($r(427) = 0.08$, $p = .112$; $F(1,408) = 1.69$, $p = .194$, respectively), so we did not include them in the following analyses.

The extent to which spelling teaching practices differed between the experimental and control groups was another important variable. As noted above, aspects of the Twictée protocol such as dictation and dictation correction are widely used components of spelling teaching in France. In addition, 14 of the 21 teachers in the control group said they conducted spelling discussions with their students.

Table 3. Features of Spelling Interventions in the Experimental and Control Groups

	Experimental group	Control group
Individual writing under dictation	+	+
Collaborative writing (spelling discussions)	+	+
Dictation correction	+	+
Collaborative identification of spelling mistakes in other students' writing	+	
Writing feedback on other students' spelling mistakes (Twictools)	+	
Categorizing spelling mistakes	+	
Collaborative editing using feedback provided by other students	+	

3.3 Research Design

Our sample comprised 40 classes in the Paris and Grenoble areas whose teachers volunteered to take part in the study during the 2017-2018 academic year. All 40 teachers had shown an interest in the way spelling is taught, were interested in the study, and were comfortable enough with their teaching practices to allow us to observe and film their classroom sessions.

3.3.1 Experimental and Control Groups

The experimental group contained 19 classes whose teachers already used Twictée and who answered a call for volunteers from the Twictée Association. These

teachers followed the Twictée protocol at their own pace throughout the school year, with no input from the researchers. Each teacher decided when and how often they used Twictée. Not providing guidelines on how frequently teachers had to use Twictée enabled us to determine whether the number of cycles influenced student performance. Each class completed between three and six Twictée cycles. The control group contained 21 classes whose teachers had attended training courses in language studies and who had responded to our call for volunteers. They received no introduction to Twictée and were asked to follow their usual spelling teaching methods ("business as usual"). We evaluated the performances of the students in both groups at the beginning (pretest) and end (posttest) of the school year and subjected the resulting data to multilevel analysis.

3.3.2 Data Collected

Each teacher completed a questionnaire asking for socio-demographic information, teaching experience, and teaching practices. The researchers also conducted semi-structured interviews with each teacher at the beginning and the end of the schoolyear in order to explore their beliefs about spelling, spelling teaching practices, and Twictée.

For each class, we videoed and transcribed either a Twictée cycle (experimental group) or five spelling lessons (control group). In the case of the Twictée cycles, we also collected students' individual dictations, group productions, tweets, and assessment dictations, mainly for qualitative analysis. We also videoed students' group work in some classes.

The third dataset included in our analyses consisted of the students' pretest and posttest spelling evaluations.

3.3.3. Participants

Classes. Our sample covered three school grades and consisted of eleven 4th-grade classes, sixteen 5th-grade classes, and four 6th-grade classes. In addition, eight classes combined 4th and 5th grades, and one class combined 3rd and 4th grades (labeled "mixed-level classes"). The distribution of grades between the experimental and control groups differed slightly. Eighteen of the classes were in schools within France's Priority Education Network (Réseau d'éducation prioritaire), which was set up in 1981 to try and close the gap in academic performance between students from privileged versus disadvantaged socioeconomic backgrounds. The Priority Education Network provides additional financial support, mostly in the form of extra teaching hours and compensatory credit, to schools in socially disadvantaged areas (assessed in terms of unemployment rate, socioeconomic status, housing quality, etc.) whose students do not reach certain academic performance criteria.

Table 4. Description of the Class, Teacher, and Student Variables. Data are N (%) or Mean (*SD*)

Classes/Teachers (N = 40)		
Condition	Experimental	19 (47.5%)
Priority education	Yes	17 (42.5%)
Class size[a]	M(sd)	24.9 (3.2)
Class size - dichotomized	[18,25]	24 (60%)
	(25,33]	16 (40%)
Mixed-level class	Yes	9 (22 %)
Teacher gender	Female	30 (75%)
	Missing	1 (2.5%)
Teaching experience[a] in years	M(*SD*)	15.4 (7.7 %)
Teaching experience in years	[2,14]	20 (50%)
	(14,37]	19 (47.5%)
	Missing	1 (2.5%)
Students (N = 893)		
Gender	Female	456 (51.1%)
	Missing	1 (0.2%)
Socioeconomic status[b]	Very disadvantaged	50 (5.6%)
	Disadvantaged	335 (37.6%)
	Intermediate	238 (26.7%)
	Advantaged	219 (24.5%)
	Missing/Unknown	50 (5.6%)
Grade	4th	355 (39.8%)
	5th	447 (50%)
	6th	91 (10.2%)
Fluency score (MCLM)	M (*SD*)	116 (40)

[a] Class Size and Teaching Experience were dichotomized before inclusion in the models.
[b] Based on the highest socioeconomic status of the two parents' professions.

Teachers. Teachers completed a questionnaire through which they provided personal data and information about their training and experience. The teachers in our sample had an average age of 42.8 years (*SD* = 7.12), which is the same as the mean age of teachers in France (MENESR, 2018, p. 271). There was no significant difference in the mean ages of the teachers in the experimental and control groups. Mean length of teaching experience was 15 years for teachers in both groups. Most of the teachers (60.5%) had a 3-year university diploma, but 36.8% of them had a 4-year university diploma or higher.

Students. The classes in our sample included a total of 997 students, 406 of whom attended Priority Education Network schools. Excluding ten students who were still in 3rd grade and all the students who were absent for the pretest and/or posttest

gave a final sample of 893 students (see Table 4). We collected several types of sociodemographic data for the students, including date of birth, gender, and mother's and father's professions. Of these 893 students, 832 (93.2%) spoke French at home (66.6% spoke French only; 26.5% spoke French and another language at home).

3.3.4. Assessments

Students completed the following three spelling tests in September 2017 (pretest) and in June 2018 (posttest).

1) A dictation used by France's Directorate for Evaluation, Forecasting, and Performance (DEPP) in its comparison of 5th-grade students' spelling abilities between 1987 and 2015 (DEPP, 2016):

Le soir tombait. Papa and maman, inquiets, se demandaient pourquoi leurs quatre garçons n'étaient pas rentrés.
- Les gamins se sont certainement perdus, dit maman. S'ils n'ont pas encore retrouvé leur chemin, nous les verrons arriver très fatigués à la maison.
- Pourquoi ne pas téléphoner à Martine? Elle les a peut-être vus !
Aussitôt dit, aussitôt fait !
À ce moment, le chien se mit à aboyer.

Night was falling. Dad and Mom, worried, were wondering why their four boys had not come home.
- The kids have surely gotten lost, Mom said. If they haven't yet found their way, they'll be very tired when they get home.
- Why don't we call Martine? She might have seen them!
No sooner said than done!
At that moment, the dog started barking.

2) A transformation exercise, which involved rewriting a text after putting the first noun (la sauterelle/grasshopper) into the plural and modifying the rest of the text accordingly:

Lis le texte suivant :
La sauterelle est un insecte nuisible. Elle adore manger les feuilles and les fleurs; elle se nourrit aussi parfois d'insectes. La sauterelle envahit souvent les plantations qu'elle est capable de saccager. L'an dernier, elle a encore ravagé des champs de blé au Brésil. Mais, demain, la sauterelle grillée arrive dans nos assiettes !

Recopie ce texte en commençant par "Les sauterelles". Modifie tout ce qu'il faut modifier.

Read the following text:
The grasshopper is a destructive insect. It loves eating leaves and flowers; sometimes it also eats insects. The grasshopper often invades crops, which it is capable of destroying. Last year, it once again ravaged Brazil's wheat fields. But, tomorrow, grilled grasshopper will be on our plates!

Copy this text but start with "The grasshoppers". Modify all the other words that need to be modified.

3) Writing a text. These results are not presented here.

The students also completed the MCLM fluency test, which measures the number of words a student can read correctly in a minute (Lequette et al., 2011). We integrated fluency into the multilevel model to analyze the relationship between students' progress in spelling and their fluency.

3.3.5. Study Variables

These tests allowed us to obtain a global score consisting of the number of words and the number of punctuation marks written/transcribed correctly in each of the dictation tests. We also calculated sub-scores for 14 variables covering the main difficulties encountered when writing French (see Table 5).

3.4 Data Analysis

We used the Factorial Analysis for Mixed Data (FAMD) principal components method to impute missing data in the covariates (see Table 4). After examining the distributions of co-variables, we dichotomized Teaching Experience and Class Size, as we did not expect to find a linear relationship between these variables and the students' performances. Due to the absence of a theoretical or empirical cut-off point, we used median values to dichotomize these variables. To simplify the statistical models, and because the modalities were unbalanced, we also recoded the four socioeconomic status categories into two modalities, labeled Disadvantaged (combining the initial Very Disadvantaged and Disadvantaged categories) and Advantaged (combining the initial Intermediate and Advantaged categories). We used median values (interquartile range) in the case of continuous variables and frequency (proportion) in the case of binary variables to describe the data (see Table 4).

Table 5. Study Variables

Variable name	Content	Items	Number of items
Verb Forms			
PP with auxiliary *to have*	Endings of past participles used with the auxiliary *avoir* (*to have* - no agreement with preceding direct object)	(ont) ravagé -(n'ont) retrouvé	2
Infinitive *-er*	Verb infinitives ending in *-er*	saccag**er** - mang**er** – aboy**er** – télépho**ner** - arriv**er**	5
3rd person plural of auxiliary verbs	3rd person plural present indicative forms of auxiliary verbs	sont - sont - ont - sont - ont	5

Say and *put* in simple past, 3rd person singular	3rd person singular simple past forms of the verbs *dire* and *mettre*	dit - mit	2
3rd person plural of verb ending not pronounced	Non-audible 3rd person plural, present indicative verb endings	ador**ent** - arriv**ent**,	2
3rd person plural of verb ending pronounced	3rd person plural present indicative verb endings	nourriss**ent** - envahiss**ent**	2
Imperfect, 3rd person singular and plural	3rd person singular and 3rd person plural imperfect tense verb endings	demand**aient** - n'ét**aient** pas rentrés - tomb**ait**	3
E verb forms	Spelling of verb forms ending in /E/	ravag**é**, mang**er**, saccag**er**, aboy**er**, retrouv**é**, téléphon**er**, aboy**er**, arriv**er**, tomb**ait**, n'ét**aient**, demand**aient**	11
Frequent verbs	3rd person singular and 3rd person plural present indicative and simple past forms of commonly used verbs	sont - sont - ont - sont - ont - dit - mit	7
Plural Markers of Nouns and Adjectives			
Subject pronoun, 3rd person plural	Plural forms of pronouns	**Elles** (adorent) - **elles** (se nourrissent) - **elles** (sont) - **elles** (ont) - s'**ils** (n'ont)	5
Plural noun	Plural forms of nouns	Les sauterelle**s** - des insecte**s** - les sauterelle**s** - les sauterelle**s** - les gamin**s** - (leurs quatre) garçon**s**	5
Adj nb agreement	Plural forms of adjectives	Nuisible**s** - capable**s** - grillée**s** - inquiet**s**	4
Lexical Spelling			
Invariable words	Spelling of invariable words	parfois, aussi, souvent - pourquoi, encore, très, aussitôt	7
Lexical spelling	Lexical writing (excluding invariable words)	soir, papa, maman, tomb-, demand-, quatre, garçon, rentr-, gamin, perdu-, retrouv-, chemin, arriv-, fatigu-, maison, téléphon-, moment	17
Global Score			
Global Score	Total number of words and punctuation marks correctly written/transcribed in each dictation tests.	All the words and punctuation marks in the two dictation tests.	138

Global Score was analyzed using linear mixed-effects models with random effects for students and classes. We began by using a model comparison strategy based on likelihood ratio chi-square statistics.

We started with an empty model (model 0), to which we successively added the main effects of time (pre- vs. post-intervention) and condition (control, experimental) as fixed effects (model 1), the time x condition interaction (model 2), the main effects of student and class characteristics as fixed effects (see Table 5) (model 3), and the Characteristics x Time x Condition three-way interaction (model 4). Because the distribution of Global Scores deviated from normality and showed a slight negative skew, we applied an ordered quantile normalization transformation before conducting further analyses.

We used the same method to analyze the 14 sub-scores, but we did not include the 4th model (three-way interactions) or the model selection method. Model 3 was the final model for each of the sub-scores (i.e., model including all the main effects of characteristics). Because the sub-scores were not normally distributed and normalization procedures did not produce satisfactory results, we dichotomized the sub-scores (using quantiles) and used generalized linear mixed-effects models with log link to analyze the resulting binomial variables.

All analyses were performed using R 4.0.5. Appendix 2 lists intra-class correlation estimates for the different models.

4. Results

4.1 Model Comparisons and Test of the Main Hypothesis

If Twictée has a positive impact on students' spelling performance, the increase in Global Scores between the pretest and posttest measures should be greater in the experimental group than in the control group. Hence, we tested Twictée's effectiveness in improving performance by looking at the Time x Condition interaction (i.e., Model 2 vs. Model 1) (see Table 6).

Table 6. Model Comparisons for Sub-Scores

	Model 1 vs. Model 0 $X^2(2)$ (p-value)		Model 2 vs. Model 1 $X^2(1)$ (p-value)		Model 3 vs. Model 2 $X^2(10)$ (p-value)	
PP with auxiliary to have	9.97	(0.007)	0.91	(0.339)	37.72	(<0.001)
Infinitive -er	3.26	(0.195)	0.08	(0.777)	51.35	(<0.001)
P6 of auxiliary verbs	81.12	(<0.001)	4.59	(0.032)	289.83	(<0.001)
Putting "say" and "put" into the simple past, 3rd person singular	57.31	(<0.001)	6.57	(0.010)	191.56	(<0.001)
P6 verb ending not pronounced	33.70	(<0.001)	0.00	(0.974)	139.86	(<0.001)

|						
P6 verb ending pronounced	77.35	(<0.001)	0.14	(0.704)	266.96	(<0.001)
Imperfect, 3rd person singular and plural	130.05	(<0.001)	0.01	(0.956)	258.96	(<0.001)
E verb forms	48.89	(<0.001)	1.18	(0.277)	229.45	(<0.001)
Frequent verbs	72.83	(<0.001)	5.75	(0.017)	262.00	(<0.001)
Subject pronoun, 3rd person plural	23.18	(<0.001)	2.49	(0.115)	199.78	(<0.001)
Plural noun	70.58	(<0.001)	1.54	(0.214)	298.26	(<0.001)
Adj-n° agreement	80.82	(<0.001)	1.11	(0.292)	173.72	(<0.001)
Invariable words	48.78	(<0.001)	1.11	(0.293)	208.11	(<0.001)
Lexical spelling	35.67	(<0.001)	6.16	(0.013)	246.02	(<0.001)

Adding the Time x Condition interaction to the Global Score model did not reveal any significant difference in improvement in spelling performance between the experimental and control groups. However, this interaction had a significant effect for 4 of the 14 sub-scores: third person plural of auxiliary verbs (χ^2 (1) = 4.59, p = 0.032, standardized estimate = 0.55), putting "say" and "put" into the simple past, third person singular (χ^2 (1) = 6.57, p = 0.010, standardized estimate = 0.60), frequent verbs, third person singular and third person plural (χ^2 (1) = 5.75, p = 0.017, standardized estimate = 0.65), and lexical spelling (χ^2 (1) = 6.16, p = 0.013, standardized estimate = 0.59).

4.2 Global Score – Final Model

Adding the Characteristics x Time x Condition interaction to the model for Global Scores (model 4) significantly improved the fit of the data (χ^2 (1) = 167.96, $p < 0.001$). Appendix 1 provides a detailed description of model 4. In this section, we focus on the Time x Condition interaction in order to evaluate Twictée's efficacy. The Time x Condition interaction in model 4 was not significant (F(1, 893) = 0.03, p = .865), nor were six of the three-way interactions: Time x Condition x Student's Gender (F(1, 893) = 0.004, p = .947); Time x Condition x SES (F(1, 893) = 0.03, p = .863); Time x Condition x School Grade (F(2, 893) = 2.74, p = .065); Time x Condition x MCLM (F(1, 893) = 1.63, p = .202); Time x Condition x Teaching Experience (F(1, 893) = 0.83, p = .363); and Time x Condition x Mixed-Level (F(1, 893) = 2.93, p = .087).

However, three of the three-way interactions were significant. The Time x Condition x Teacher's Gender interaction (Figure 2E) suggests that Twictée was more beneficial in classes with female teachers (F(1, 893) = 6.85, p = .009, partial η^2 = 0.006).

Figure 2. Marginal Normalized Mean Global Scores According to Time x Condition x Covariate Interactions.
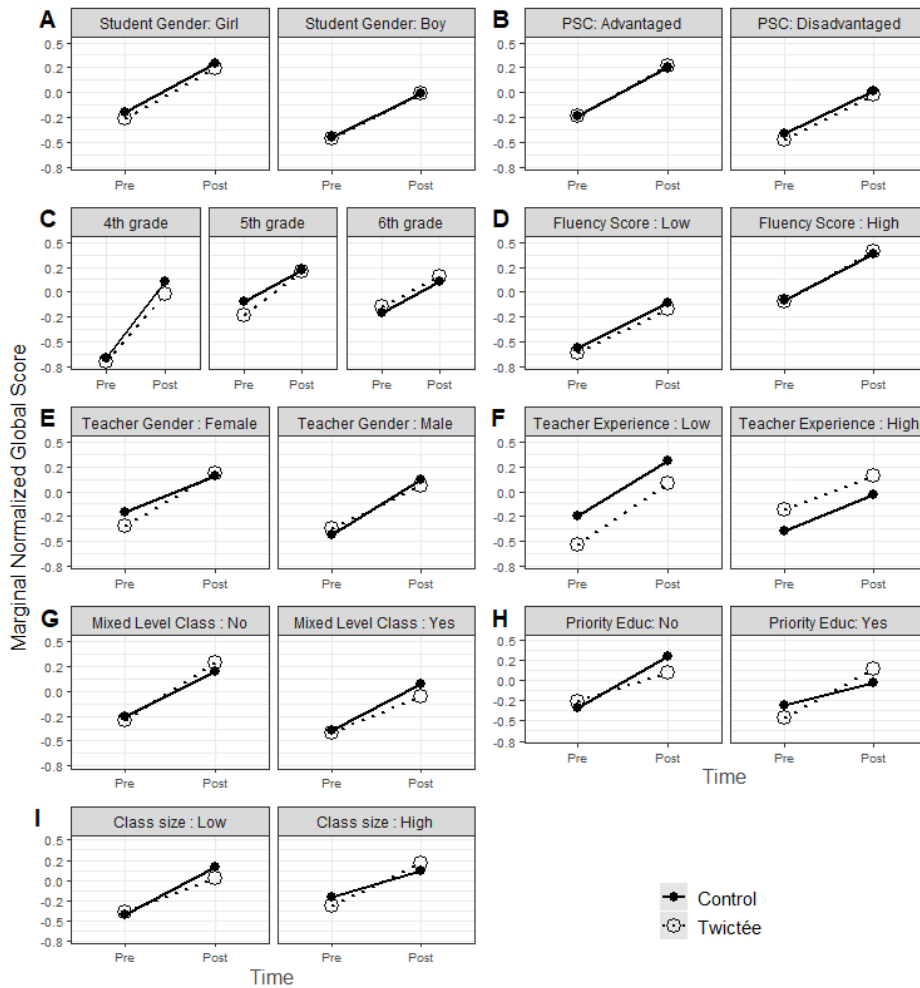
The Time x Condition x Priority Education interaction (Figure 2H) shows that in the case of Priority Education schools, students in the control group performed better on the pretest but students in the experimental group performed better on the posttest (F(1, 893) = 17.83, p < .001, partial η^2 = 0.016). Thus, students at Priority Education schools appear to have benefitted significantly more from Twictée than did students at non-Priority Education schools. Finally, the Time x Condition x Class Size interaction (F(1, 893) = 10.02, p = .002, partial η^2 = 0.009) suggests that Twictée

was less beneficial than traditional teaching methods in the case of small classes but more beneficial than traditional teaching methods in the case of large classes.

4.3 Sub-Scores – Final Models

Appendix 3 provides details of the results for the 14 sub-scores. After controlling for individual and class characteristics, the Time x Condition interaction remained significant for only four of the sub-scores: third person plural of auxiliary verbs, putting "say" and "put" into the simple past, third person singular, frequent verbs in the third person singular and plural, and lexical spelling. Improvement on these four sub-scores was greater for students in the experimental group than it was for students in the control group (OR = 1.80, 95CI = [1.07, 3.03], p =0.028; OR = 1.83, 95CI = [1.14, 2.91], p =0.011; OR = 2.03, 95CI = [1.16, 3.56], p =0.014; OR = 1.80, 95CI = [1.12, 2.90], p = .015, respectively).

5. Discussion

International studies have shown the benefits of writing under dictation, collaborative writing and editing, and developing spelling consciousness through metacognitive reflection. However, only a few studies have examined spelling in French, a language whose morpho-syntactic characteristics make spelling difficult, and none of these studies investigated the effect of combining all these aspects in a single intervention. The current study addressed this research gap by assessing the efficacy of Twictée, a French intervention that promotes metacognitive discussions of language by combining writing under dictation with collaborative preparation of feedback and collaborative editing using other students' feedback to correct mistakes.

Our study's main limitation arises from its design, which did not allow us to measure randomized controlled effects. However, all the participating teachers had shown interest in the way spelling is taught and a desire to teach it actively and positively, either by taking part in Twictée Association activities or by attending relevant training courses. They were also sufficiently interested in the study and comfortable enough with their teaching practices and classroom management to allow us to observe and film their classroom sessions. These factors suggest that any differences in motivation, engagement, and classroom management between teachers in the experimental and control groups were likely to be negligible.

Second, although our implementation measure included the number of Twictée steps each class performed, the amount of time dedicated to each Twictée cycle, and the number of cycles completed, we did not assess the quality of Twictée sessions. This methodological limitation must be taken into account when interpreting our results, as a qualitative study of four classes in our sample showed substantial variability in the quality of teaching interventions, especially in terms of

teacher-student interactions during classroom discussions (Brissaud et al., 2019; Crinon & Viriot-Goeldel, 2021).

The Global Score results showed that, on average, the students made significant progress during the school year, but students who used Twictée did not progress significantly more than control group students (see Appendix 1). Results for the sub-scores for different types of difficulty in French spelling showed significant interaction effects in the case of simple spellings, such as the third person plural forms of the auxiliary verbs *être* (*sont*) and *avoir* (*ont*) and the spellings of frequently used verbs (*sont* and *ont*, again, plus *dire* and *mettre*, which are often used in the third person singular in the simple past). However, there were no significant differences between the experimental and control groups in students' abilities to apply grammatical agreement procedures or to differentiate between verbs ending in -er, and, for example, past participles ending in -é.[3] Twictée's designers and the teachers who use it believe that Twictée helps students overcome complex spelling problems by encouraging them to think more analytically about spelling (Crinon & Viriot-Goeldel, 2021). Nevertheless, the only significant differences we found in favor of Twictée were for words that are quite simple to spell and memorize, rather than for those that are more difficult and require grammatical reasoning, and evidence for the robustness of these effects is not very strong.

These results could be due, at least in part, to variations in the way teachers implemented Twictée. All the teachers followed the first two Twictée steps and the vast majority followed steps 3, 4, and 5. The only step that a large number of teachers did not follow was the optional assessment step (step 6). In contrast, results showed differences in implementation intensity, measured via either the number of Twictée cycles followed during the school year or the amount of time allocated to Twictée cycles. Twictée is not a rigid intervention, so teachers are free to either apply all its main steps or, as many teachers do, to customize one or more of these steps as they see fit. Similarly, teachers are not required to spend set amounts of time on each step and may lengthen or shorten any step as they wish.

Teachers in the experimental group highlighted the importance of Twictée's collaborative discussions, carried out with the whole class and in small groups, during which students explain their spelling choices (Crinon & Viriot-Goeldel, 2021). However, 14 of the 19 teachers in the control group also included collaborative discussions in their spelling lessons, following a procedure known as *dictée négociée* (negotiated dictation). This finding was unexpected because *dictée négociée* is a relatively new approach and not mentioned in the national curriculum. The use of *dictée négociée* by control group teachers both highlights the difficulty of controlling control groups' practices and complicates the interpretation of our results. Indeed, it means that the main difference between our experimental and control groups was in Twictée students being required to

produce feedback and to categorize spelling mistakes, and this difference may not have been sufficient to have a significant impact on our results.

Twictée's failure to help students acquire spellings requiring complex grammatical reasoning may be due to other differences in how Twictée is implemented in the classroom that our fidelity and intensity indicators did not capture. We addressed this question by conducting a series of more detailed analyses of verbal interactions within the classroom, which may not always be good enough to foster learning. For the results of these analyses, see Brissaud et al. (2019), Fenoglio (2019), and Fenoglio and Brissaud (2020). This consideration led us to examine the conditions under which student interactions, in the whole class and in small groups, can promote collaborative learning. In terms of classroom discussions, the time allocated to student discussions in the classroom (Twictée does not include any time constraints) and the quality of these discussions varied greatly between groups. Poor anticipation of a text's difficulty and insufficient knowledge of the spelling system were sometimes the root of the unproductive (or even counterproductive) exchanges we saw during several sessions. The grammatical reasoning conducted in whole-class correction sessions is often unclear and does little to help students understand French spelling and identify the knowledge they have to apply to determine a correct spelling (Brissaud et al., 2019).

These interactions show that methods such as Twictée cannot overcome training deficits in teaching spelling and/or in managing student interactions. Preliminary analyses of interactions within groups of students have highlighted several recurring phenomena, such as poor explanatory interactions. Because Twictée relies greatly on exchanges between peers, if teachers do not provide adequate scaffolding or guidance, responsibility for teaching is, in some ways, transferred to the students. Input from the teacher is also vital when it comes to choosing the correct hashtag to categorize the spelling mistakes the other students have to correct, as students in our sample frequently chose a hashtag at random (Fenoglio, 2019). This could explain why some of the Twictée classes were among the most effective in our sample, in that their students made more progress than others, whereas other Twictée classes were among the least effective. These contrasting results show how difficult it is to measure the impact of methods that place few constraints on teachers and that allow a lot of time for interactions, but which do not include training for teachers and students on how to approach these interactions. Moreover, reflections on the value and limitations of using such methods in schools must go beyond simply evaluating their effectiveness in boosting performance.

Although a few targeted training sessions have been run for teachers in France (e.g., two quantitative training sessions on plural markers, Weth, Ugen, Fayol & Bîlici, 2021), only one intervention involving collaborative spelling discussions and mobilizing metacognitive reasoning has been shown to have positive effects on

several dimensions of French orthographics (Fisher & Nadeau, 2014). Despite integrating these elements and others, such as writing spelling feedback and categorizing spelling mistakes, Twictée failed to improve overall spelling performance. However, we found two interesting three-way interactions that suggest possible avenues for future research on Twictée. First, the fact that Twictée significantly improved spelling performance in large classes could be due to the effects of collaborative writing and editing, which give students in large classes more opportunities to discuss spelling and how language functions than they would otherwise have had in classes of this size. Second, it would be interesting to determine why Twictée had a greater impact on spelling performance for students in Priority Education schools than for students in other schools. Given the enthusiasm Twictée has aroused in educational circles, we intend to explore this aspect further by specifically studying Twictée's impact in schools in disadvantaged areas.

The present study also indicated a few methodological improvements—including randomizing the experiment, and monitoring control group teaching practices, especially spelling discussions—that we will incorporate in future investigations. Because the videos of the sessions we recorded showed differences in the quality of discussions which usual implementation measures would not capture, future studies will also include qualitative analyses. These studies will enable us to provide the method's designers and the Twictée Association with further recommendations for improving Twictée and how it is used, in addition to the recommendations arising from the current study, most notably the need for more extensive teacher training.

Endnotes

[1] Teachers also collaborate to produce the text for each Twictée and to oversee its implementation. This aspect will be investigated in further papers.

[2] Percentages based on the 38 teachers who answered this question.

[3] When hearing a French verb ending in [E], students must analyze the sentence in order to determine whether it should be written "é" (past participle) or "er" (infinitive).

Acknowledgments

References

Arabyan, M. (1990). La dictée dialoguée [Dialogued Dictation]. *L'École des Lettres Collège, 12,* 59-80.

Block, K. K., & Peskowitz, N. B. (1990). Metacognition in spelling: Using writing and reading to self-check spellings. *The Elementary School Journal, 91*(2), 151–164. https://doi.org/10.1086/461643

Bosse, M.-L., Brissaud, C., & Le Levier, H. (2020). French Pupils' Lexical and Grammatical Spelling from Sixth to Ninth Grade: A Longitudinal Study. *Language and Speech*. https://doi.org/10.1177/0023830920935558

Brissaud, C., & Cogis, D. (2011). *Comment enseigner l'orthographe aujourd'hui ?* [*How should spelling be taught today?*]. Paris: Hatier.

Brissaud, C., Cogis, D., & Totereau, C. (2014). *La performance orthographique à l'articulation école collège: une approche qualitative des marques du pluriel* [*Spelling Abilities at the Transition from Elementary School to Middle School: A Qualitative Approach to Plural Markers*]. Paper presented at the 4th World Conference on French Linguistics (CMLF), July 2014, Berlin, Germany. https://doi.org/10.1051/shsconf/20140801375

Brissaud, C., Viriot-Goeldel, C., & Ponton, C. (2019). Enseigner et apprendre l'orthographe avec la « Twictée ». Premiers résultats de l'évaluation d'un dispositif innovant d'enseignement de l'orthographe [Teaching and learning spelling with Twitter. First results of the Twictée programme]. *Repères, 60,* 107-130. https://doi.org/10.4000/reperes.2346

Bouwer, R., & De Smedt, F. (2018). Introduction Special Issue: Considerations and Recommendations for Reporting on Writing Interventions in Research Publications. *Journal of Writing Research, 10*(2), 115-137. https://doi.org/10.17239/jowr-2018.10.01.01

Chervel, A. (2006). *Histoire de l'enseignement du français du XVIIe au XXe siècle* [History of the Teaching of French from the 17th to the 20th Century]. Paris: Retz.

Cogis, D., Fisher, C., & Nadeau, M. (2015). Quand la dictée devient un dispositif d'apprentissage [When Dictation Becomes a Teaching Method]. *Glottopol, 26,* 69-91.

Cordewener, K., Hasselman, F., Verhoeven, L., & Bosman, A. (2018). The Role of Instruction for Spelling Performance and Spelling Consciousness. *The Journal of Experimental Education, 86*(2), 135-153, https://doi.org/10.1080/00220973.2017.1315711

Crinon, J., & Viriot-Goeldel, C. (2021). Intérêt et limites de la notion de dispositif [Value and limits of the French concept of "dispositif". The example of a spelling intervention]. *Pratiques, 189-190*. https://doi.org/10.4000/pratiques.10198

Direction de l'évaluation et de la prospective et de la performance [DEPP] (1996). Connaissances en français et en calcul des étudiants des années 20 et d'aujourd'hui [Students' Abilities in French and Arithmetic from the 1920s to Today]. *Note d'information,* 96.19.

Direction de l'évaluation, de la prospective et de la performance [DEPP] (2016). Les performances en orthographe des étudiants en fin d'école primaire (1987-2007-2015) [Students' Spelling Performances at the End of Elementary School (1987-2007-2015)]. *Note d'information,* 28.

Durlak, J. A., & DuPré, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3), 327-350. https://doi.org/10.1007/s10464-008-9165-0

Fenoglio, P. (2019). Le hashtag pour catégoriser les erreurs d'orthographe au cycle 3: un hiatus entre objectifs pédagogiques et appropriations des étudiants [Using the Hashtag to Categorize Spelling Mistakes in 4th to 6th Grade: A Gap Between Teaching Objectives and its Appropriation by Students]. *Repères, 60,* 131-151. https://doi.org/10.4000/reperes.2351

Fenoglio, P., & Brissaud, C. (2020). Entre tâche et activité : analyse des échanges langagiers d'élèves de cycle 3 lors de l'utilisation d'un outil didactique de correction orthographique

[Between task and activity: grade 4-6 students' verbal interactions when using a didactic tool aimed at correcting spelling mistakes], *Recherches en Éducation*, *40*, 10-27. https://doi.org/10.4000/ree.427

Fisher, C., & Nadeau, M. (2014). Usage du métalangage et des manipulations syntaxiques au cours de dictées innovantes dans des classes du primaire [The Use of Metalanguage and Syntactic Manipulations in the Course of Innovative Dictations in Primary School Classes]. *Repères*, *49*, 169-191. Doi: https://doi.org/10.4000/reperes.742

Galuschka, K., Görgen, R., Kalmar, J., Haberstroh, S., Schmalz, X., & Schulte-Körne, G. (2020). Effectiveness of spelling interventions for learners with dyslexia: A meta-analysis and systematic review. *Educational Psychologist*, *55*(1), 1-20. https://doi.org/10.1080/00461520.2019.1659794

Geist, B., Kupetz, M., & Glaser, K. (2019). Accounting for spellings: Ambisyllabic consonants in L2 German spelling discussions. *Classroom Discourse*, *10*(1), 71-98. https://doi.org/10.1080/19463014.2019.1567361

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*(4), 879–896. https://doi.org/10.1037/a0029185

Graham, S. & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445-476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing*, *27* (1), 1703-1743. https://doi.org/10.1007/s11145-014-9517-0

Hobart, F., & Forgione, R. (2015). *Twictée*. Retrieved from the Twictée association website http://www.twictee.org/twictee

Jaffré, J.-P. (1992). Le traitement élémentaire de l'orthographe. Les procédures graphiques [Basic Processing of Spelling. Graphic Procedures]. *Langue française*, *95*, 27-48. https://doi.org/10.3406/lfr.1992.5770

Jaffré, J.-P. (2004). Peut-on parler de sémiographie optimale ? [Is There Such a Thing as Optimal Semiography?]. *LIDIL*, *30*, 11-25. https://doi.org/10.4000/lidil.593

Kreiner, D. S., & Green, A. G. (2000). Feeling-of-knowing judgments and strategy selection in spelling. *Perceptual and Motor Skills, 90*, 775–786. https://doi.org/10.2466/pms.2000.90.3.775

Lequette, C., Pouget, G., & Zorman, M. (2011). *Fluence* [*Fluency*]. Grenoble: La Cigale Editions.

Manesse, D. & Cogis, D. (2007). *Orthographe: à qui la faute ?* [*Spelling: Whose Mistake is it?*]. Paris: ESF.

Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche (MENESR) (2018). Programme d'enseignement du cycle des apprentissages fondamentaux (cycle 2) [Teaching Program for Foundational Skills (1st to 3rd grade)]. *Journal Officiel*, n° 30 of July 26, 2018, Retrieved from https://www.education.gouv.fr/pid285/bulletin officiel.html?cid bo=132987

Morin, M.-F., Alamargot, D., Diallo, T., & Fayol, M. (2018). Individual differences in lexical and grammar spelling across primary school. *Learning and Individual Differences*, *62*, 128-140. https://doi.org/10.1016/j.lindif.2018.02.002

Paveau, M. A. (2016). L'écriture numérique. Standardisation, delinéarisation, augmentation [Digital Writing. Standardization, Delinearization, Augmentation]. *Fragmentum*, *48*, 13-36. https://doi.org/10.5902/fragmentum.v0i48.23296

Peereman, R. (1999). LEXOP: A Lexical Database Providing Orthography-Phonology Statistics for French Monosyllabic Words. Behavior Research Methods. *Instruments and Computers*, *31*(2), 376-379. https://doi.org/10.3758/BF03207735

Resnicow K., Davis, M., Smith, M., Lazarus-Yaroch, A., Baranowski, J., & Wang, D.T. (1998). How to best measure implementation of school health curricula: A comparison of three measures. *Health Education Research*, *13*(2), 239-250. https://doi.org/10.1093/her/13.2.239

Riou, J. (2017). Quels sont les effets du temps passé à encoder sur les performances des élèves en décodage et en orthographe au cours préparatoire ? [The effects of time spent encoding on student performance]. *Repères*, *55*, 159-182. https://doi.org/10.4000/reperes.1163

Sprenger-Charolles L. (2004). Linguistic Processes in Reading and Spelling: The Case of Alphabetic Writing Systems: English, French, German and Spanish. In: T. Nunes, P. Bryant (Eds), *Handbook of Children's Literacy* (pp. 43-65). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-1731-1_4

Totereau, C., Thevenin, M., & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to Spell* (pp. 97-114). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vanderswalmen, R., Vrijders, J., & Desoete, A. (2010). Metacognition and Spelling Performance in College Students (pp. 367-394). In A. Efklides & P. Misailidi (Eds.), *Trends and Prospects in Metacognition Research*. New York: Springer. https://doi.org/10.1007/978-1-4419-6546-2_17

Viriot-Goeldel, C. et Brissaud, C. (2019). Enseigner et apprendre l'orthographe aujourd'hui [Teaching and learning spelling]. *Repères*, *60*, 15-22. https://doi.org/10.4000/reperes.2271

Weth, C., Ugen, S., Fayol, M., & Bîlici, N. (2021). Spelling patterns of plural marking and learning trajectories in French taught as a foreign language. *Written Language and Litteracy, 24*(1), 81-109. https://doi.org/10.1075/wll.00048.wet

Wiliams, K., Walker, M., Vaughn, S., & Wanzek, J. (2017). A Synthesis of Reading and Spelling Interventions and Their Effects on Spelling Outcomes for Students With Learning Disabilities. *Journal of Learning Disabilities*, *50*(3) 286-297. https://doi.org/10.1177/0022219415619753.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical Analysis of the Bidirectional Inconsistency of Spelling and Sound in French. Behavior Research Methods. *Instruments and Computers*, *28*(4), 504-515. https://doi.org/10.3758/BF03200539

Appendix 1: Model 4 Results for Global Score

Parameter	Coefficient	95% CI	F value	P-value
Time [post]	0.66	[0.44, 0.89]	122.525	< 0.001
Condition [experimental]	-0.12	[-0.63, 0.40]	0.045	0.833
Student Gender [male]	-0.25	[-0.37, -0.14]	43.707	< 0.001
SES [disadvantaged]	-0.18	[-0.32, -0.04]	28.153	< 0.001
Level [4th grade]	-0.46	[-0.83, -0.09]	15.556	< 0.001
Level [5th grade]	0.11	[-0.24, 0.46]		
Fluency Score (MCLM)	0.48	[0.41, 0.55]	528.887	< 0.001
Teacher Gender [male]	-0.23	[-0.53, 0.07]	1.361	0.251
Teaching Experience (14.37)	-0.15	[-0.35, 0.05]	0.028	0.867
Priority Education [yes]	-0.13	[-0.46, 0.20]	2.861	0.099
Mixed-Level Class [yes]	0.02	[-0.33, 0.37]	1.047	0.312
Class Size (25.33)	0.21	[-0.06, 0.48]	1.476	0.231
Time x Student Gender	-0.06	[-0.16, 0.04]	2.237	0.135
Condition x Student Gender	0.05	[-0.12, 0.22]	0.452	0.502
Time x Condition	-0.18	[-0.50, 0.14]	0.029	0.865
Time x SES	-0.06	[-0.17, 0.06]	1.505	0.220
Condition x SES	-0.06	[-0.25, 0.13]	0.366	0.545
Time x Level [4th grade]	0.46	[0.23, 0.70]	30.935	< 0.001
Time x Level [5th grade]	0.01	[-0.21, 0.22]		
Condition x Level [4th grade]	-0.10	[-0.61, 0.42]	0.171	0.843
Condition x Level [5th grade]	-0.20	[-0.68, 0.29]		
Time x Fluency Score (MCLM)	0.01	[-0.04, 0.07]	3.724	0.054
Condition x Fluency Score (MCLM)	0.04	[-0.06, 0.13]	1.960	0.162
Time x Teacher Gender	0.20	[0.02, 0.38]	0.976	0.324
Condition x Teacher Gender	0.19	[-0.18, 0.56]	0.064	0.801
Time x Teaching Experience	-0.20	[-0.32, -0.07]	22.316	< 0.001
Condition x Teaching Experience	0.51	[0.18, 0.85]	8.245	0.007
Time x Mixed-Level Class	0.00	[-0.19, 0.20]	2.711	0.100
Condition x Mixed-Level Class	0.01	[-0.43, 0.45]	0.242	0.625
Time x Priority Education	-0.35	[-0.56, -0.13]	0.532	0.466
Condition x Priority Education	-0.21	[-0.67, 0.24]	0.140	0.710
Time x Class Size	-0.26	[-0.42, -0.09]	1.849	0.174
Condition x Class Size	-0.13	[-0.50, 0.24]	0.078	0.781
Time x Condition x Student Gender	0.01	[-0.14, 0.15]	0.004	0.947
Time x Condition x SES	0.01	[-0.15, 0.18]	0.030	0.863
Time x Condition x Level [4th grade]	-0.09	[-0.40, 0.23]	2.738	0.065
Time x Condition x Level [5th grade]	0.12	[-0.16, 0.42]		
Time x Condition x Fluency Score - MCLM	0.05	[-0.03, 0.14]	1.629	0.202
Time x Condition x Teacher Gender	-0.29	[-0.51, -0.07]	6.854	0.009
Time x Condition x Teaching Experience	-0.09	[-0.29, 0.11]	0.827	0.363
Time x Condition x Mixed-Level Class	-0.23	[-0.49, 0.03]	2.929	0.087
Time x Condition x Priority Education	0.59	[0.32, 0.87]	17.827	< 0.001
Time x Condition x Class Size	0.36	[0.14, 0.59]	10.017	0.002

Appendix 2: Intra-class Correlations

	Student	Class
Global score	0.421	0.275
PP with auxiliary *to have* PRE	0.133	0.037
Infinitive *-er* PRE	0.117	0.091
P6 of auxiliary verbs PRE	0.191	0.235
Say and *put* in simple past 3rd person singular PRE	0.183	0.137
P6 verb ending not pronounced PRE	0.261	0.244
P6 verb ending pronounced PRE	0.302	0.363
Imperfect 3rd person singular and plural PRE	0.194	0.194
E verb forms PRE	0.218	0.135
Frequent verbs 3rd person singular and plural PRE	0.211	0.226
Subject pronoun 3rd person plural PRE	0.178	0.128
Plural noun PRE	0.197	0.168
Adj nb spelling PRE	0.117	0.085
Invariable words PRE	0.181	0.135
Lexical spelling PRE	0.256	0.111

Appendix 3: Results of the Generalized Linear Mixed-Effects Models with Log Links for the 14 Sub-Scores

Top panel

	P6¹ of auxiliary verbs		Frequent verbs		PP with auxiliary *to have*		Imperfect P3 P6	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Time [post]	2.60	[1.80, 3.74] ***	2.53	[1.72, 3.73] ***	1.35	[0.94, 1.93]	5.29	[3.48, 8.05] ***
Condition [experimental]	0.76	[0.46, 1.23]	0.74	[0.44, 1.26]	0.86	[0.55, 1.35]	1.05	[0.57, 1.94]
Student Gender [male]	0.56	[0.41, 0.76] ***	0.43	[0.30, 0.61] ***	0.84	[0.63, 1.12]	0.61	[0.44, 0.85] **
SES [disadvantaged]	0.47	[0.33, 0.67] ***	0.45	[0.30, 0.66] ***	0.97	[0.71, 1.34]	0.44	[0.30, 0.65] ***
Level [4th]	0.21	[0.10, 0.47] ***	0.21	[0.09, 0.47] ***	1.38	[0.69, 2.76]	0.36	[0.13, 0.97] *
Level [5th]	0.48	[0.24, 0.96] *	0.46	[0.23, 0.93] *	1.64	[0.89, 3.03]	1.01	[0.41, 2.47]
Fluency (MCLM)	3.31	[2.65, 4.13] ***	3.30	[2.58, 4.23] ***	1.44	[1.22, 1.70] ***	3.50	[2.73, 4.49] ***
Teacher Gender [male]	0.63	[0.38, 1.05]	0.60	[0.35, 1.04]	0.60	[0.38, 0.93] *	1.11	[0.58, 2.13]
Teaching Experience (14.37)	1.21	[0.75, 1.94]	1.45	[0.88, 2.40]	0.96	[0.64, 1.44]	0.65	[0.35, 1.22]
Priority Education [yes]	0.99	[0.58, 1.69]	0.91	[0.51, 1.61]	1.27	[0.79, 2.03]	0.68	[0.34, 1.37]
Mixed-Level Class [yes]	0.57	[0.35, 0.93] *	0.72	[0.43, 1.21]	1.04	[0.70, 1.55]	0.62	[0.33, 1.17]
Class Size (25.33)	1.36	[0.78, 2.39]	0.99	[0.55, 1.78]	1.48	[0.90, 2.42]	1.11	[0.54, 2.30]
Time x Condition	1.80	[1.07, 3.03] *	2.03	[1.16, 3.56] *	1.30	[0.77, 2.19]	0.99	[0.57, 1.71]

Bottom panel

	Say and put in simple past P3		Lexical spelling		E verb forms		Invariable words	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Time [post]	1.87	[1.36, 2.57] ***	1.57	[1.13, 2.18] **	2.06	[1.49, 2.85] ***	2.92	[2.02, 4.21] ***
Condition [experimental]	0.69	[0.46, 1.04]	0.72	[0.48, 1.06]	0.78	[0.50, 1.22]	0.93	[0.58, 1.49]
Student Gender [male]	0.52	[0.39, 0.68] ***	0.64	[0.48, 0.85] **	0.65	[0.50, 0.86] **	0.59	[0.44, 0.79] ***
SES [disadvantaged]	0.79	[0.58, 1.07]	0.73	[0.53, 1.00] t	0.55	[0.40, 0.75] ***	0.82	[0.59, 1.14]
Level [4th]	0.50	[0.25, 1.01] t	0.66	[0.36, 1.23]	0.81	[0.39, 1.68]	0.73	[0.36, 1.45]
Level [5th]	0.70	[0.37, 1.33]	1.24	[0.72, 2.13]	1.03	[0.53, 2.00]	1.06	[0.58, 1.96]
Fluency (MCLM)	2.34	[1.96, 2.80] ***	2.93	[2.40, 3.57] ***	2.76	[2.28, 3.33] ***	2.81	[2.30, 3.44] ***
Teacher Gender [male]	0.74	[0.49, 1.12]	1.00	[0.68, 1.47]	0.86	[0.54, 1.36]	1.09	[0.68, 1.74]
Teaching Experience (14.37)	1.25	[0.84, 1.88]	0.88	[0.61, 1.27]	1.02	[0.65, 1.58]	1.16	[0.75, 1.79]
Priority Education [yes]	0.69	[0.43, 1.10]	0.96	[0.63, 1.46]	0.95	[0.57, 1.56]	0.76	[0.46, 1.25]
Mixed-Level Class [yes]	1.14	[0.76, 1.70]	0.65	[0.45, 0.95] *	0.71	[0.45, 1.11]	0.70	[0.45, 1.10]
Class Size (25.33)	0.66	[0.40, 1.09]	1.41	[0.90, 2.18]	1.00	[0.59, 1.70]	1.10	[0.66, 1.82]
Time x Condition	1.83	[1.15, 2.91] *	1.81	[1.12, 2.90] *	1.28	[0.80, 2.03]	0.74	[0.44, 1.23]

Notes. P-values were recoded as: t for p < 0.07; * for p < 0.05; ** for p < 0.01 and *** for p < 0.001

---

¹ P6 refers to the 3rd person plural and P3 refers to the 3rd person singular.

Appendix 4: Results of the Generalized Linear Mixed-Effects Models with Log Links for the 14 Sub-Scores

	P6 verb ending pronounced			Plural noun			Infinitive -er			Subject pronoun P6		
	OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI	
Time [post]	2.15	[1.50, 3.08]	***	3.34	[2.35, 4.74]	***	1.17	[0.87, 1.56]		1.51	[1.09, 2.10]	*
Condition [experimental]	0.55	[0.29, 1.05]		0.94	[0.60, 1.50]		0.89	[0.59, 1.36]		0.86	[0.59, 1.26]	
Student Gender [male]	0.54	[0.39, 0.76]	***	0.40	[0.30, 0.53]	***	0.94	[0.74, 1.18]		0.64	[0.49, 0.84]	**
SES [disadvantaged]	0.59	[0.40, 0.87]	**	0.65	[0.48, 0.89]	**	1.09	[0.84, 1.43]		0.67	[0.49, 0.91]	*
Level [4th]	0.24	[0.08, 0.70]	**	0.42	[0.20, 0.91]	*	1.24	[0.61, 2.52]		0.22	[0.12, 0.39]	***
Level [5th]	0.62	[0.24, 1.65]		1.10	[0.56, 2.18]		1.59	[0.83, 3.04]		0.39	[0.24, 0.63]	***
Fluency (MCLM)	2.31	[1.86, 2.87]	***	2.75	[2.28, 3.32]	***	1.50	[1.30, 1.72]	***	2.23	[1.88, 2.66]	***
Teacher Gender [male]	0.79	[0.39, 1.61]		0.67	[0.41, 1.08]		0.82	[0.53, 1.28]		1.24	[0.88, 1.76]	
Teaching Experience (14.37)	1.61	[0.82, 3.16]		0.91	[0.57, 1.44]		0.75	[0.48, 1.15]		1.22	[0.87, 1.70]	
Priority Education [yes]	1.35	[0.63, 2.89]		0.70	[0.41, 1.18]		1.26	[0.77, 2.06]		1.10	[0.74, 1.62]	
Mixed-Level Class [yes]	0.81	[0.41, 1.61]		0.74	[0.46, 1.17]		0.68	[0.44, 1.06]		0.91	[0.64, 1.29]	
Class Size (25.33)	1.36	[0.61, 3.02]		1.30	[0.75, 2.27]		2.24	[1.33, 3.79]	**	1.13	[0.76, 1.68]	
Time x Condition	0.99	[0.58, 1.67]		0.73	[0.45, 1.18]		1.06	[0.70, 1.61]		1.46	[0.91, 2.34]	

	P6 verb ending not pronounced			Adj nb agreement		
	OR	95% CI		OR	95% CI	
Time [post]	3.88	[2.50, 6.02]	***	2.51	[1.84, 3.44]	***
Condition [experimental]	0.53	[0.22, 1.27]		0.72	[0.51, 1.03]	
Student Gender [male]	0.35	[0.23, 0.55]	***	0.67	[0.52, 0.86]	**
SES [disadvantaged]	0.43	[0.26, 0.70]	**	0.64	[0.48, 0.85]	**
Level [4th]	0.08	[0.02, 0.40]	**	0.66	[0.40, 1.10]	
Level [5th]	0.40	[0.09, 1.77]		1.27	[0.82, 1.98]	
Fluency (MCLM)	4.68	[3.36, 6.53]	***	1.90	[1.63, 2.22]	***
Teacher Gender [Male]	0.69	[0.25, 1.92]		0.60	[0.43, 0.83]	**
Teaching Experience (14.37)	1.23	[0.46, 3.29]		0.91	[0.67, 1.23]	
Priority Education [yes]	1.06	[0.35, 3.19]		1.04	[0.73, 1.49]	
Mixed-Level Class [yes]	0.33	[0.12, 0.89]	*	0.72	[0.52, 0.98]	*
Class Size (25.33)	2.07	[0.64, 6.67]		1.10	[0.77, 1.59]	
Time x Condition	1.08	[0.60, 1.97]		1.28	[0.81, 2.00]	

Notes. P-values were recoded as: [t] for $p < 0.07$; * for $p < 0.05$; ** for $p < 0.01$ and *** for $p < 0.001$

Appendix 5. Raw means for the Time x Condition interaction

	Twictée				No Twictée			
	Pre		Post		Pre		Post	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Global score	48,07	15,05	55,49	13,64	49,86	13,98	56,45	12,37
PP with auxiliary *to have* PRE	0,77	0,67	0,85	0,73	0,82	0,68	0,88	0,71
Infinitive *-er* PRE	2,93	1,27	2,99	1,19	2,89	1,29	3,08	1,20
P6 of auxiliary verbs PRE	3,07	1,72	3,85	1,45	3,26	1,64	3,83	1,41
Say and *put* in simple past 3rd person singular PRE	1,29	0,75	1,59	0,63	1,41	0,70	1,56	0,64
P6 verb ending not pronounced PRE	0,74	0,81	0,94	0,84	0,84	0,84	1,10	0,82
P6 verb ending pronounced PRE	0,97	0,95	1,26	0,91	1,08	0,95	1,39	0,85
Imperfect 3rd person singular and plural PRE	0,84	1,01	1,37	1,22	0,83	1,00	1,36	1,19
E verb forms PRE	5,09	2,13	5,95	2,52	5,09	2,08	6,01	2,41
Frequent verbs 3rd person singular and plural PRE	4,37	2,20	5,45	1,82	4,67	2,02	5,39	1,74
Subject pronoun 3rd person plural PRE	3,32	1,65	3,86	1,37	3,41	1,63	3,89	1,25
Plural noun PRE	4,37	1,70	4,93	1,42	4,44	1,71	5,15	1,24
Adj nb spelling PRE	1,57	1,21	2,10	1,28	1,68	1,27	2,29	1,13
Invariable words PRE	5,15	1,47	5,61	1,30	5,25	1,39	5,75	1,28
Lexical spelling PRE	13,58	3,22	14,73	2,41	14,17	2,76	14,76	2,21