Navigating the Challenges of L2 Reading: Self-Efficacy, Self-Regulatory Reading Strategies, and Learner Profiles

SUZANNE GRAHAM,[1] ROBERT WOORE,[2] ALISON PORTER,[3] LOUISE COURTNEY,[4] AND CLARE SAVORY[5]

[1] *University of Reading, Institute of Education, London Road Campus, 4 Redlands Road, Reading, RG1 5EX, United Kingdom Email: s.j.graham@reading.ac.uk*

[2] *University of Oxford, Department of Education, 15 Norham Gardens, Oxford, OX2 6PY, United Kingdom Email: robert.woore@education.ox.ac.uk*

[3] *University of Southampton, Department of Modern Languages and Linguistics, Avenue Campus, Southampton, SO17 1BF, United Kingdom Email: amp1g09@soton.ac.uk*

[4] *University of Reading, Institute of Education, London Road Campus, 4 Redlands Road, Reading, RG1 5EX, United Kingdom Email: l.m.courtney@reading.ac.uk*

[5] *University of Oxford, Department of Education, 15 Norham Gardens, Oxford, OX2 6PY, United Kingdom Email: clare.savory@education.ox.ac.uk*

Reading in a foreign language has value for learners as a potentially rich source of input as well as enjoyment. It requires persistence, however. Within models of self-regulated learning, persistence relates to learners' self-efficacy and use of strategies to aid task completion and regulation of engagement. Yet the relationship between self-efficacy and self-regulatory strategies is underexplored for second language (L2) reading, despite some intervention studies finding that instruction aimed at improving strategy use positively influences self-efficacy. The current study investigated the relationship between what we call text engagement regulatory reading strategies (TERRS) and reading self-efficacy among 529 beginner learners of French. It also explored whether different learner profiles exist with respect to that relationship, and how far learners of different profiles benefited in respect of reading self-efficacy from 3 instructional approaches: phonics instruction plus the use of challenging texts; strategy-based instruction using the same texts; and no explicit phonics or strategy instruction using the texts only. The use of TERRS was an important predictor of reading self-efficacy and central to 3 distinct learner profile clusters. Increases in reading self-efficacy were significantly greater for learners of certain profiles who received strategy-based instruction, with implications for theories of self-regulated language learning and classroom practice.

Keywords: French; reading; self-efficacy; self-regulation; strategies

The Modern Language Journal, 0, 0, (2020)
DOI: 10.1111/modl.12670
0026-7902/20/1–22 $1.50/0

THIS STUDY INVESTIGATED SELF-EFFICACY for reading among young learners of French as a foreign language in England, that is, their belief in their ability to complete French reading activities successfully. In particular, this study explored the relationship between reading strategies and reading self-efficacy, including whether different learner profiles exist with respect to this relationship. It also considered the impact on reading self-efficacy for different learner profiles of three

different types of reading instruction: (a) phonics-focussed instruction using challenging reading texts, (b) reading strategies and self-regulation-focussed instruction using the same challenging reading texts, or (c) no explicit phonics or strategy instruction when reading the same challenging texts.

An investigation into the self-efficacy of young language learners is timely and important because there is growing evidence that the earlier start to instructed language learning that has become a global trend (Murphy, 2014) does not guarantee high levels of self-confidence amongst learners (Graham et al., 2016). Learning a foreign language in the classroom is a long and often arduous endeavour, particularly in instructed contexts where there is little input from the second language (L2) in the natural environment, where instructional hours are low, and where there are few instrumental reasons to acquire the language. These characteristics pertain especially to learning a foreign language in England, but are also common to other Anglophone contexts such as Australia (Kohler, 2017).

Also common to those contexts are relatively low levels of motivation for and persistence with foreign language learning; for example, in England, fewer than 50% of high-school learners choose to continue with language study beyond the compulsory phase (13–14 years of age). According to some researchers, learners' lack of persistence in pursuing language study in England is attributable, at least in part, to low levels of self-efficacy and a poor sense of progress in language learning (e.g., Erler & Macaro, 2011), even among those whose level of achievement is high when measured by national examination results (Graham, 2004). This last point suggests that a focus on raising learning outcomes alone is not sufficient to encourage learners to continue language study. Furthermore, self-efficacy may be especially fragile and unstable when learners move from primary (elementary) to secondary (high) school, perhaps explaining why studies have found that a decline in motivation for language learning begins in the first year of secondary school (Courtney, 2017; Graham et al., 2016). In early adolescence, learners start to make more normative comparisons between themselves and peers that can lead them to judge that they are less proficient in a curriculum area than they thought (Wigfield et al., 1997), particularly when faced with changes in instructional emphases or approaches (Courtney, 2017). Variability in primary language provision in England also means that at the start of secondary school, teachers are faced with learners of wide-ranging language proficiency, leading to difficulties in ensuring appropriate progress for them all (Graham et al., 2016). In such situations, learners may revise previous self-efficacy judgements (Linnenbrink & Pintrich, 2003), potentially in a downward direction.

SELF-EFFICACY AND SELF-REGULATORY STRATEGIES

While self-efficacy can be fragile for young language learners, it is also central, determining "how much effort students will expend on an activity, how long they will persevere when confronting obstacles, and how resilient they will be in the face of adverse situations" (Pajares, 2002, p. 116). Similarly, studying college learners of English in South Korea, Yun, Hiver, and Al–Hoorie (2018) found that students' ability to cope with such setbacks (what they termed 'buoyancy') was strongly predicted by both self-efficacy and self-regulated learning (SRL) strategy use.

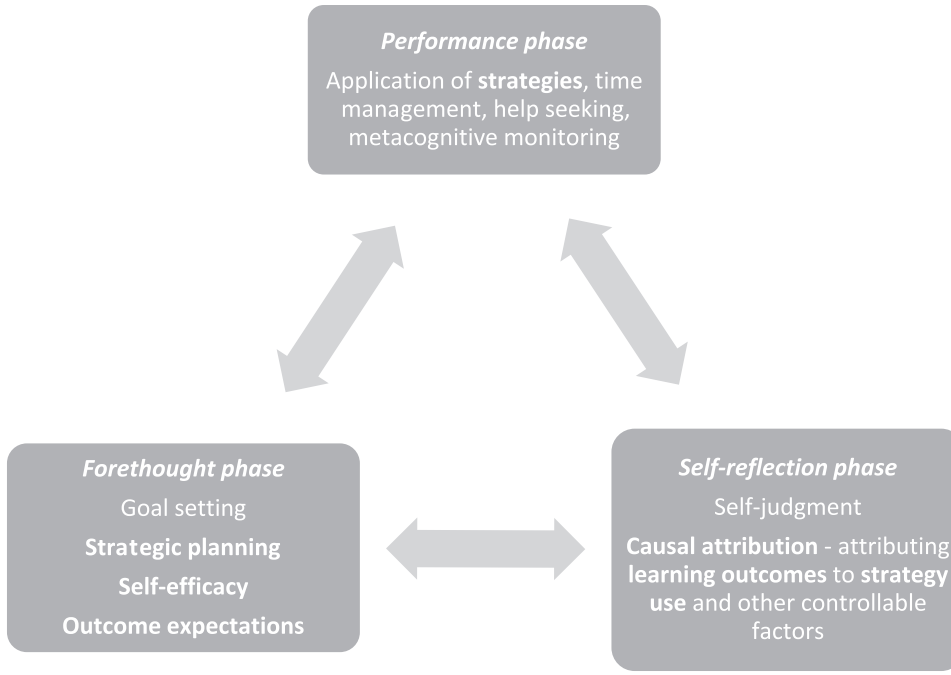
The relationship between self-efficacy and self-regulatory strategies has been explored relatively infrequently for L2 learning, however, especially among young, beginner learners. Reviewing 32 articles published between 2003 and 2012 that considered self-efficacy and language learning, Raoofi, Tan, and Chan (2012) identified only 6 that explored this relationship directly and a further 7 that did so indirectly by investigating improvements in self-efficacy after strategy-based interventions. All 13 studies identified a positive relationship between self-efficacy and self-regulatory strategies, but almost none were with school-aged learners, leaving underexplored a group of learners for whom low self-efficacy, and hence lack of persistence for language study, is a concern.

Definitions and Models of Self-Regulation

The relationship between self-efficacy and self-regulatory strategies and how they both feed into persistence is alluded to in definitions of SRL by L2 researchers. For instance, Kormos and Csizér (2014) observed that SRL constitutes "self-regulatory control that involves the use of strategies which are largely conscious processes that students apply to control their learning" (p. 279). Indeed, self-efficacy and strategic behaviour come together in several models of self-regulation beyond the L2 field, particularly in those based on social cognitive theory (Panadero, 2017).

One of the most influential and empirically supported models (Panadero, 2017) is that of

FIGURE 1
Self-Regulated Learning



Note. Adapted from Zimmerman (2013).

Zimmerman (2013). Although undergoing various modifications, at its core are three cyclical stages, in which self-efficacy and strategy use come together at all points, but particularly in the crucial first and last phases. The first, the forethought phase, involves analysis of the learning task, goal setting, and plan formation to address those goals, including strategy selection; and the activation of a range of motivational beliefs including self-efficacy. In the second, performance phase, learners monitor their progress during task execution, apply strategies to facilitate task completion, and regulate their engagement and motivation. Self-observation and self-recording are important aspects of this second phase, during which learners undertake "tracking of specific aspects of their own performance, the conditions that surround it, and the effects that it produces" (Zimmerman, 2000, p. 19). They also then feed into the final, self-reflection phase, where learners assess their performance and offer explanations for the task outcome, in other words, make causal attributions. Higher levels of self-efficacy for future tasks are most likely to occur if self-observation and self-recording have enabled learners to understand the positive connection between how they carried out the task and the level of success they achieved, thus helping them grasp that they are the originators of their own achievements. Learners who attribute outcomes to their own strategies—that is, to "correctable causes" (Zimmerman, 2000, p. 23)—are likely to experience a greater sense of control over learning and personal agency during the acquisition of self-regulation (Graham & Macaro, 2008), potentially leading to greater self-efficacy, persistence, and hence potentially better learning outcomes. A successful learning outcome, especially on a challenging task, will also help in the development of 'mastery experiences' (Bandura, 1997). These in turn impact on self-efficacy and sense of competence to undertake similar tasks again in the future, and so the cycle begins again. This implies that there is a reciprocal relationship between self-efficacy, strategy use, and learning outcomes, linked by causal attributions, as indicated in bold in Figure 1, showing an adapted version of Zimmerman's (2013) cycle.

Instructional Frameworks for Self-Regulated Learning

The cycle outlined in Figure 1 has influenced frameworks for SRL instruction such as Zimmerman's (2013) multilevel model, in which learners are taught to use strategies as part of

SRL. Instruction typically involves, first, observation of the teacher, who models the strategy or strategies in question. Learners then emulate what has been observed through deliberate and focussed practice (coupled with support and feedback from the teacher and/or peers and input on how and when to use which strategy). More autonomous execution of the task and self-regulatory use of the taught strategies then follow, in which the learner aims to select the most effective strategies for the task at hand. Self-observation and metacognitive reflection are important aspects of such instructional models, whereby learners are encouraged to monitor and evaluate the effectiveness of the strategies they are using. This also encourages them to make strategy-based attributions for learning outcomes, which, as outlined previously, contribute to self-efficacy development. Although infrequently referring explicitly to SRL frameworks, L2 strategy instruction intervention studies—especially more recent ones—draw implicitly upon them and include similar steps (see Gu, 2019). As such, they follow what Ardasheva et al. (2017) called an "awareness-raising" model (p. 554), following four steps from awareness raising, through modelling and practice that gradually becomes less guided, to self-evaluation, strategy evaluation, and transfer of strategies to new contexts (Gu, 2019).

Self-Efficacy, Self-Regulation, and L2 Reading

While self-efficacy and self-regulation are important for all aspects of language learning, they may be particularly so for the complex task of reading, especially at the beginner stage (Graham et al., 2016), and even more so where linguistic development is slow because of limited exposure to the L2 inside or outside the classroom—as is the case in England. Being able to read in the foreign language potentially has value for learners, not only as a rich source of input but also from a motivational perspective. Confident and self-regulated beginner learners could, in the longer term, engage in extensive reading, believed to support young learners' L2 reading motivation (Briggs & Walter, 2013) and L2 reading proficiency (Jeon & Day, 2016). Reading a range of texts is a normal part of most high-school learners' everyday experience in using their first language (L1), both as a source of information and for interest and enjoyment. Inspection of coursebooks in England suggests, however, that learners have limited exposure to reading material of that kind in the L2 (Andon & Wingate, 2013), even if one allows for the relatively small vocabulary sizes reported for learners in that context (Milton, 2006). If learners never feel able to access interesting and enjoyable texts in the L2—even after several months or even years of instruction—then arguably their sense of being able to communicate effectively, and hence their sense of self-efficacy as a foreign language learner, may be impaired.

Furthermore, becoming literate in the foreign language poses challenges because of the complexity of L2 reading, involving a combination (Grabe & Jiang, 2013) of lower level processes (e.g., word recognition and phonological decoding) and higher level processes (e.g., drawing on schematic knowledge to infer meaning and monitoring comprehension). The application of self-regulatory reading strategies forms part of those higher level processes, which contribute not only to reading outcomes per se but also to ensuring that learners engage in enough reading to become more proficient at it rather than giving up in the face of difficulty. These strategies thus help learners to "engage in reading, to expend effort, to persist in reading without distraction" (Grabe & Jiang, 2013. 4) and hence perhaps undertake more extensive reading, with the possibility of longer term benefits for reading proficiency.

Individual Differences and Self-Regulated Learning Interventions

Within meta-analyses across both L1 and L2 contexts, moderate to large effect sizes are reported for strategy instruction programmes (or interventions) targeting reading comprehension, but the effect sizes vary according to a number of group and individual differences (IDs). The latter are difficult to define but we follow Ardasheva & Tretter (2013), citing Dörnyei's (2005) broad definition, viewing IDs as characteristics "on which people vary by degree" (Ardasheva & Tretter, 2013, p. 325). The IDs most commonly considered in relation to SRL are the age and the L1–L2 proficiency and linguistic knowledge (LK) of the learners involved, as well as educational level, a group difference. In L1 reviews, larger effect sizes are associated with reading strategy interventions conducted at secondary rather than at primary school level (Dignath & Büttner, 2008) and with learners who have already been taught to decode (National Reading Panel, 2000). Similarly, Taylor (2014) argued that for comprehension strategy instruction to be effective for L2 reading, learners need to have automated the lower level processes before they can allocate processing space to self-regulation and metacognition as part of strategy instruction.

This suggests that explorations of the relationship between self-efficacy and self-regulatory strategy use within the context of beginning L2 reading need to take account of issues of LK, particularly vocabulary knowledge and proficiency in phonological decoding, as important factors in early L2 reading (Grabe & Jiang, 2013). The importance of LK may also explain why reviews of the impact of L2 interventions targeting self-regulatory aspects such as self-efficacy or strategy use have found greater impact among older learners—namely, those over 12 years of age—compared with younger learners (Ardasheva et al., 2017). Younger age typically coincides with lower levels of L2 proficiency, especially in input-poor instructional contexts like England. It should be emphasised, however, that across L2 reading strategy interventions, few have been conducted with younger learners, and fewer still have considered their impact on aspects of self-regulation. It is therefore difficult to assess how beneficial L2 reading strategy interventions might be for them.

One of the few studies to consider the effectiveness of reading strategy instruction for young, beginner L2 learners and its impact on self-regulatory variables is by Macaro and Erler (2008), who in fact reported positive effects. We focus on that study because it involved the population of concern in our own study—namely, learners of French aged 11–12 in England. It investigated the impact of a reading strategy intervention on strategy use and confidence in reading as well as on reading comprehension. A group of 62 learners received explicit reading strategy instruction, while a comparison group of 54 learners followed their usual French syllabus. The intervention included strategies designed to help learners understand individual problem words and monitor their understanding of the text (e.g., sound out the word or phrase, guess the meaning of a problem word from surrounding words, read the whole sentence to see if it makes sense). To help with the sounding-out strategy, learners also received instruction in symbol–sound correspondences (SSCs). Although not articulated as such, the intervention followed many aspects of Zimmerman's (2013) multilevel model of SRL instruction, including its emphasis on scaffolding. Key aspects were (a) teacher modelling of strategies, (b) guided practice or emulation, (c) reflection on which strategies to use in which circumstances and evaluation of their effectiveness, and (d) feedback from the researchers that drew participants' attention to the link between their strategy use and their reading outcomes. The intervention group outperformed the comparison group on a challenging reading comprehension task at the end of the intervention, with a large effect size. Additionally, evidence suggested at posttest that their self-efficacy for and persistence in reading had improved. They were more willing than the comparison group to expend effort to tackle the posttest text in the first place, showed increased awareness of the importance of strategic behaviour by being more likely to use so-called text engagement strategies, and reported greater positivity towards learning French in general and reading French in particular, as assessed through a questionnaire.

The extent to which the strategy instruction as opposed to the phonics instruction was responsible for the improvements in reading and attitudes was not explored by Macaro & Erler (2008); however, the researchers argued that it was learning how to "orchestrate" strategies (p. 114), rather than changing their use of individual strategies, that helped learners. That orchestration, they continued, was facilitated chiefly through the feedback on their strategy use. This form of "verbal persuasion" (Bandura, 1997, p. 101) is also an important factor in helping learners change their thinking about task performance and the reasons behind it, with positive implications for self-efficacy (Macaro, 2019).

The study by Macaro & Erler (2008) therefore suggests that lower proficiency learners can benefit from instruction in self-regulatory reading strategies. This is also the view of Taylor (2014): Despite arguing that beginner L2 learners need first and foremost vocabulary-enhancing instruction, he still sees a place for reading strategies alongside that instruction, because of their ability to "empower the first year L2 learner" (p. 61). In other words, strategy instruction can contribute to reading development through the potential impact on self-efficacy for reading. This may be particularly true for another group of learners infrequently considered in L2 reading research: young L2 learners whose lower levels of L1 literacy may mean they have fewer L1 comprehension strategies to transfer to L2 reading (Grabe & Jiang, 2018). Their general academic performance (GAP) can also modulate strategy use (Ardasheva & Tretter, 2013). Learners with lower L1 literacy and GAP are precisely those who, in the context of England, are likely to give up foreign language study or to be withdrawn from it because it is deemed too difficult for them (Tinsley, 2019).

It is also likely that different learner characteristics can combine in complex ways, modulating the effect of SRL interventions and

also the relationship between strategy use and self-efficacy. Research has principally targeted the latter rather than the former. For example, a study of ESL learners in Botswana found that on the one hand, strategy use was greater among higher proficiency learners, and among primary school learners compared with secondary- and tertiary-level learners (Magogwe & Oliver, 2007). On the other hand, the correlation between total strategy use and self-efficacy for primary school learners was found to be stronger as proficiency level decreased.

Nor is the relationship between strategy use and other variables such as language proficiency always a linear one (Fung & Macaro, 2019). Thus, analyses other than correlation might shed greater light on how different IDs come together in different groups of learners. For example, cluster analysis is useful where "there is evidence to suggest that different subgroups of learners may utilize different pathways to language learning" (Staples & Biber, 2015, p. 244). Using that approach in a study of listening comprehension, Fung and Macaro (2019), for example, found that among 646 high-school English-as-a-foreign-language learners in Hong Kong, there was a cluster of learners with low levels of LK (vocabulary and grammar) but with levels of listening-strategy use comparable to that of some learners with high LK. The authors concluded that some learners compensate for low LK by being more 'strategic.' Similarly, Yun et al. (2018), as well as finding that both self-efficacy and self-regulatory strategy use predicted language learners' levels of buoyancy, also found that those variables were important within five learner profiles that they established through cluster analysis. For example, at one end of the spectrum was the *thriver* group with very high self-efficacy and strategic self-regulation, low levels of anxiety, and hence high buoyancy levels; at the other the *dependent* group with the very lowest levels of self-efficacy and strategic self-regulation but high anxiety and low buoyancy. Furthermore, in a subsequent structural equation modelling analysis, the importance of buoyancy for learning outcomes was not diminished when prior L2 achievement was taken into account, suggesting that self-efficacy and self-regulatory strategies, as well as being related, help explain how different groups of learners make differing use of the prior attainment they bring to the classroom. In other words, some learners seem better able than others to make the most of what skill and knowledge they have. Furthermore, identifying different learner profile groups might be helpful for informing classroom practice, particularly if one also explores the extent to which they are impacted differently by different forms of instruction—an area which, to our knowledge, has not been previously explored in L2 reading research among beginner learners.

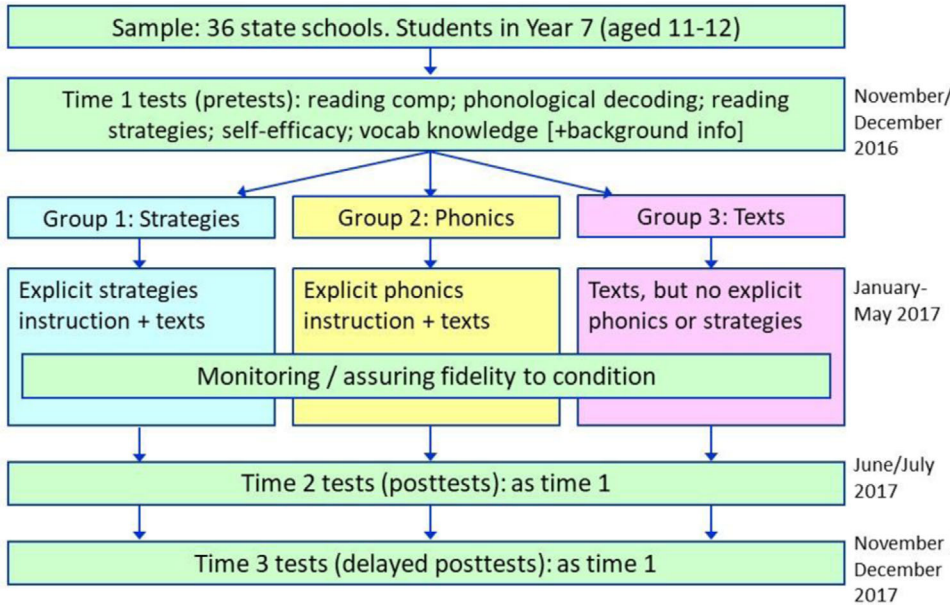
In summary, while interest in L2 learner strategy research continues—especially from an SRL perspective (Rose et al., 2018)—there is a recognised need for greater understanding of the relationship between strategies and self-efficacy, how this might vary for different types of learners, and the relative benefits of strategy instruction for learners of different profiles (Ardasheva et al., 2017; Plonsky, 2019; Zhang, 2008). In other words, more research is needed into who benefits from language learner strategy instruction, how, and why. Such a focus is particularly important at the transition from primary to secondary school language learning in contexts such as England, where variability in provision at primary school often means that language teachers encounter cohorts of learners who, at the start of secondary education, vary widely on several important variables—namely, L2 reading proficiency, LK, and GAP.

RESEARCH QUESTIONS

In light of the previous discussion, the present study aimed to explore the relationship between self-regulatory strategies and self-efficacy within the context of L2 reading among beginner learners of French, taking into account L2 proficiency (LK, L2 reading comprehension) and L1 GAP. Furthermore, we wished to respond to the call to consider language learner strategy instruction in relation to a range of learner differences (Pawlak, 2019; Plonsky, 2019), by examining the impact of a reading strategy-based intervention on self-efficacy for learners of different profiles, compared with other types of reading instruction. As such, the study addressed the following research questions:

RQ1. What is the relationship between L2 reading self-efficacy and L2 text engagement regulatory reading strategies (TERRS)?
RQ2. To what extent do different learner profiles exist with respect to this relationship?
RQ3. What is the impact of a strategy-based reading intervention on reading self-efficacy for learners with different profiles, compared with other types of reading instruction?

FIGURE 2
Study Design [Color figure can be viewed at wileyonlinelibrary.com]



METHODOLOGY

Study Design

This article is based on data from a larger study (Woore et al., 2018), a randomized control trial (RCT) assessing the impact of three different types of reading instruction, and more specifically, a cluster RCT, in which whole classes of students participated (each class with a different teacher, in different schools). The initial design was a pretest–posttest–delayed-posttest study with three intervention groups (strategies, phonics, and texts groups; see Figure 2). Because of considerable attrition at Time 3,[1] we concentrate here on data from Time 1 (T1; winter 2016) and Time 2 (T2; summer 2017). The key findings of Woore et al. (2018) were as follows: First, no group outperformed the others on T2 reading comprehension at a statistically significant level, once school factors were controlled for. However, all groups showed statistically significant T1–T2 increases in reading comprehension (and descriptively, the phonics group made the most progress). Interview and questionnaire data suggested that progress was promoted by participants' engagement with the challenging texts used with all groups. Second, for phonological decoding, the phonics group showed a descriptive advantage over the other groups at T2, with a statistically significant advantage over the texts group after controlling for general academic attainment (i.e., GAP). Third, the phonics group recorded significantly higher vocabulary knowledge at T2 compared to the texts group, and after controlling for GAP, both the strategies and phonics groups outperformed the texts group. All three groups showed significant T1–T2 increases in vocabulary scores with varying effect sizes (texts: small; strategies: small–medium; phonics: medium–large). Fourth, for self-efficacy, the strategies and phonics groups showed greater T1–T2 improvements, descriptively, than the texts group. Furthermore, after controlling for GAP, the strategies group displayed significantly higher levels of self-efficacy than the texts group at T2. Changes in total strategy use were small, although greater for the strategies group and for individual strategies. As a RCT, Woore et al. (2018) focussed on between-groups differences for the whole sample rather than examining the impact of the intervention on different groups of learners, which is the concern of the present analyses.

Context and Participants

Students were learning French as an L2 in Year 7, the first year of secondary school, age 11–12. They would have received varying amounts of instruction in French in previous school years, because of variability in the amount of

TABLE 1
Participants by Intervention Group

Intervention Group	Male	Female	Not Reported/Other	Total
Strategies	78	70	15	163
Phonics	86	114	14	214
Texts	79	72	1	152

curriculum time devoted to language study in primary schools, and in primary French teachers' proficiency levels and language pedagogy training (Tinsley, 2019). Using convenience sampling, we recruited 36 state-funded schools from England and Wales. These included 3 schools with a selective, attainment-based entry policy (*grammar* schools) and 33 nonselective, *comprehensive* schools, with one class of French learners in each school taking part (878 learners overall). Schools came from a range of locations and demographic and attainment profiles. They were allocated to intervention groups through 'minimization' (Torgerson & Torgerson, 2013), using freely available software (https://sourceforge.net/projects/minimpy/) to allocate the first school randomly to one of the three groups, and then subsequently allocate schools in such a way as to create groups which were as well-matched as possible on school type (grammar or comprehensive), and percentages of students in each school who (a) were eligible for free school meals (a proxy measure of socioeconomic disadvantage), (b) were recorded as having English as an additional language, and (c) in the year preceding the study, had achieved five or more of the top three grades in the GCSE (national examination taken at 16), including in mathematics and English. Learners who reported in a questionnaire that they spoke French at home were removed from the dataset for the analyses, but those who spoke another, non-French language other than English were retained. Not all learners completed all tests at all time points, and data on students' GAP were not provided by all schools. We therefore selected a final sample of 529 learners from 28 schools for whom valid T1 data were available for the following variables[2]: self-regulatory L2 reading strategies, L2 LK, L2 reading comprehension, and L1 GAP (see Table 1).

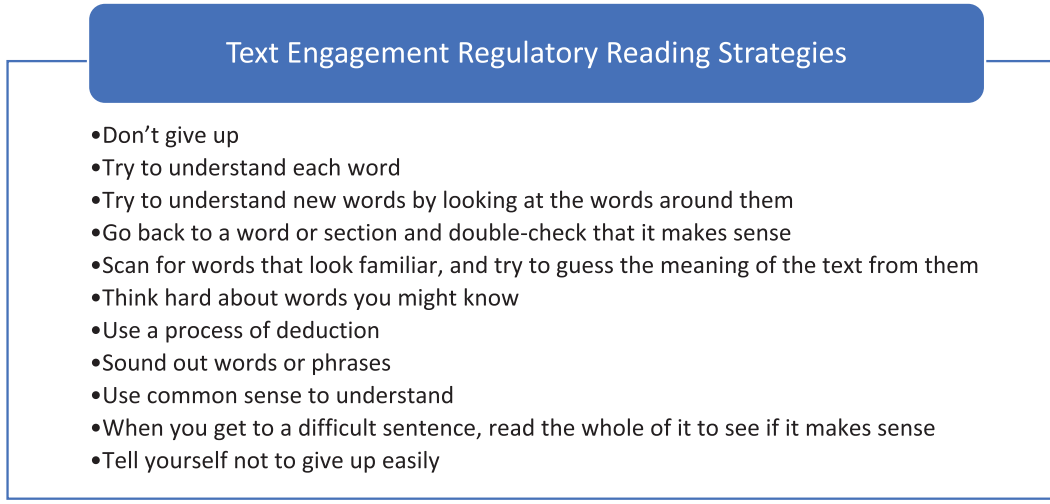
Procedures and Methods

Figure 2 outlines the time points for pre- and posttesting. Instruments used at both time points to assess each outcome variable of interest for the present study are detailed in the following sections. They were administered in the order in which they are presented. All instruments appear in the Online Supporting Information.

Reading Strategies. Reading strategy use was measured prior to the main battery of tests, which took up one whole lesson, using a self-report questionnaire adapted from Macaro & Erler (2008). Their questionnaire was "scenario based" and asked learners to assess "reading-related cognitive strategies and more general approaches" (p. 102) they used for reading in French. We modified it, however, so that it included all eight strategies taught in our intervention, giving 18 items in total plus an *other* option, not included in analyses. We ensured that the wording in the questionnaire differed from that used in the intervention materials for the strategies group (to avoid a 'teaching-to-the-test' effect). Although the original questionnaire asked learners to rate items on a 3-point scale (*yes*, *no*, *sometimes*), piloting for the present study suggested that learners found it conceptually easier to respond to items on a scale running from *never* to *very often*. Using such labels has attracted some criticism in the literature, on the grounds that frequency of strategy use should not be taken to reflect effective use (Pawlak, 2019). While acknowledging such an argument, we were guided by what, in piloting, learners seemed most able to respond to as a questionnaire item.

For analysis, reverse coding was undertaken in SPSS (version 25) so that a higher score corresponded to greater self-regulatory behaviour. Principal components analysis with direct Oblimin rotation was then applied to T1 strategies. All 18 items correlated at least .3 with at least one other item, suggesting reasonable factorability, except for one: "look up many words in dictionary" (later excluded from analysis). The Kaiser–Meyer–Olkin measure of sampling adequacy was .84, above the commonly recommended value of .6, and Bartlett's test of sphericity was significant: χ^2 (153) = 2,045.72,

FIGURE 3
The Text Engagement Factor Emerging From the Factor Analysis [Color figure can be viewed at wileyonlinelibrary.com]



$p < .001$. There was levelling off of eigenvalues on the scree plot after two factors, which we termed (a) TERRS, comprising 11 items (see Figure 3), and (b) reliance on others or avoidance strategies, comprising 6 items. These factors explained 22.47% and 13.74% of the variance, respectively. Internal consistency (Cronbach's alpha) for each of the scales ranged from high (TERRS, .82) to low–medium (reliance on others/avoidance strategies, .65). In view of the lower internal consistency of the latter scale, we based our analysis on the strategies for the TERRS scale and computed a score for each learner, summing scores for each of the 11 items (maximum total of 44). While summing strategies in this way may be criticised on the grounds that it again implies that frequency rather than manner of strategy use is the main focus (Pawlak, 2019), we felt it was justifiable as a reflection of overall text engagement, following the approach used by Macaro & Erler (2008).

Reading Comprehension. The main testing session began with reading comprehension. In the absence of a national standardized French reading test for beginners, we based the tests used at T1 and T2 closely on those designed by Macaro & Erler (2008) and, like them, included two short translation tasks and two short 'main ideas' tasks. The translation tasks assessed understanding at the level of lexis and syntax whilst the main-ideas tasks examined global understanding. One of the translation tasks and both main-ideas tasks were common across T1 and 2. The second translation task used at T2 was more difficult, to take account of learners' likely increase in French proficiency by then. As we wanted to gain direct insight into students' progress across the two time points, in the present analyses we omitted this second task and used only scores for the tasks common to both time points. The risk of a practice effect was considered to be low, as 6 months elapsed between the two rounds of testing and no feedback was given after T1.

The content, grammar, and vocabulary in the translation and main-ideas tasks were consistent with the Year 7 French curriculum, but at a high enough level of difficulty to avoid a ceiling effect as ascertained through piloting. For analysis, a scoring scheme for the translation was drawn up by the research team by determining units of meaning. One point was allocated for each unit of meaning, giving a maximum score of 29. For the two main-ideas tasks, the main ideas in each text were decided upon through discussion by the team, giving a maximum score of 16 across the two texts. Both sets of scores were then combined to give a 'common items' score at T1 and T2 (out of 45). The reliability of the scoring scheme (for all reading tasks) was then assessed through two rounds of pilot scoring of around 10% of T1 scripts, with modifications made after the first round of any aspects of the scheme that did not seem to function well. An interrater reliability rate of .98 for the translation task and .96 for the main-ideas tasks was then obtained for the revised scoring scheme as implemented by two scorers in the second round. Any remaining differences in

scores were discussed and salient points noted on the scoring scheme for use in the final scoring of all scripts.

Self-Efficacy for Reading French. We measured reading self-efficacy using self-report questionnaire items adapted from Graham & Macaro (2008). The questionnaire was completed immediately after the reading comprehension tasks and as such asked learners to reflect on their level of confidence in being able to complete similar tasks in the future. It asked about very specific aspects of L2 reading that reflected what learners would normally do in class as well as key aspects of the intervention: the ability to read such a text all the way to the end, to understand the main ideas and details it contained, to work out the meaning of unknown words, and to read the text aloud. Cronbach's alpha was .82 at T1 and .84 at T2. We therefore calculated T1 and T2 composite self-efficacy scores for each learner by summing individual item responses.

Phonological Decoding. This was measured using a pen-and-paper 'sound-alike task' (SALT) developed specifically for the study and used rather than an individual reading aloud test, which the large sample size precluded. In the SALT, each of the items presented participants with a set of three French pseudowords. Using knowledge of sublexical SSCs, learners had to decide which (if any) of the three pseudowords sounded the same. Pseudowords were created by taking real words from the 1,000 most frequent French words (based on Lonsdale & Le Bras, 2009) and changing the onset consonant(s), checking with two native speakers that no resulting pseudoword was a real French word. All words were monosyllabic and had roughly equal numbers of letters. One point was awarded for each pair of words correctly identified as sounding the same. Piloting showed acceptable reliability for the test (Cronbach's alpha of .85), which also correlated significantly with a reading aloud test (as a criterion validity measure) conducted with a sample of 32 participants not involved in the main study, $r = .85$, $p < .001$. In the final dataset in the main study, Cronbach's alpha was somewhat lower, but still acceptable, being .73 at both time points.

Vocabulary Knowledge. Breadth of vocabulary knowledge was assessed using a modified, pen-and-paper version of X-Lex (Milton, 2006), a self-report test assessing the most basic form of vocabulary knowledge: vocabulary recognition (i.e., participants are asked to indicate whether they know a series of French words). To increase the validity of the test, a number of pseudowords are included, allowing a correction to be applied for guesswork. We reduced the length of the test from 120 items to 48, including items from only the 2,000 most frequent French words (rather than the 5,000 most frequent words in the original), in keeping with the low proficiency levels of our learners. We drew 20 words from the 1,000 most frequent words[3] and 20 words from the 2,000 most frequent words, plus 8 pseudowords (thus maintaining the same proportion of real words to pseudowords as in the original test).

Learners' estimated vocabulary size was calculated as the number of correctly identified real words multiplied by 50 minus the number of incorrectly ticked pseudowords multiplied by 250 (following David, 2008). We removed any students with a 'false alarm rate' (i.e., incorrect identification of pseudowords as real words) greater than 20%, that is, three or more false alarms (J. Milton, personal communication, March 12, 2018), to maximize test reliability. In the few instances where negative scores remained, we corrected these to 0 to indicate that the students concerned had demonstrated no measurable knowledge on the test.

Scores for vocabulary and phonological decoding were converted to z scores and then combined to give a total LK score at T1 and T2. We used this combined measure of LK because, for the current article, our focus was more on the relationship between self-efficacy and TERRS when LK as a whole was accounted for rather than on the impact of our interventions on different aspects of LK, which was the focus of the original study (Woore et al., 2018). We also followed the approach of Fung & Macaro (2019) in using a composite measure of LK, albeit substituting phonological decoding for grammatical knowledge, given the importance of decoding for early L2 reading (Grabe & Jiang, 2013).

L1 General Academic Performance. Schools provided scores for learners from standardized tests assessing verbal and mathematical performance, administered at either the end of primary school or the start of secondary school. As schools used different test versions, we converted all scores to z scores.

Interventions: The Programmes of Instruction

The interventions were delivered by class teachers over a period of 16 weeks, for 20–30 minutes per week for each of the three groups, with care taken to ensure equal amounts of exposure to the

same reading material across all groups. Like most 11–12 year olds in England (Tinsley, 2019), in total our learners received approximately 2 hours of language instruction per week. Each group worked with a series of eight challenging texts (using language above learners' current productive level) created especially for the study and accompanied by questions and a link to a public, online video clip related to the text's theme. The questions were mainly simple comprehension questions designed to be easily answerable, for example, by recognizing numerals or clear cognates; they were supplemented with a small number of more open-ended, reflective questions. The research team worked with participating teachers to co-create intervention materials (see examples in the Online Supporting Information) and lesson plans for the teachers to implement in their classrooms. Teachers received approximately 6 hours of training in using the materials and the principles underpinning them. They also agreed to refrain from giving explicit instruction pertaining to the other groups for the duration of the intervention. Fidelity to condition was monitored and judged acceptable through lesson observations and weekly teacher logs of all reading-related instruction throughout the intervention.

Strategies. The strategy-based intervention followed the principles implemented by Macaro & Erler (2008), as well as drawing on Zimmerman's (2013) multilevel model of SRL instruction. Initial tasks were designed to raise learners' awareness of the kinds of strategies they could use to help them understand challenging texts. They were then given a checklist (available in both English and French) containing eight strategies that were introduced and revisited throughout the intervention. The focus was on self-regulation of comprehension and text engagement, not only suggesting certain ways of working out a word's meaning (e.g., looking at the words surrounding the 'problem' word), but also emphasizing the need for perseverance and for evaluating any decisions made to see if comprehension had been achieved.

Subsequently, learners worked with the eight texts to practise using the strategies. Most sessions began with some explicit modelling of strategy use by the teacher, taking a short portion of the text to show how the eight strategies could be used to help learners understand, and how they might be combined. Learners were then asked to tackle additional paragraphs and identify the main points of information conveyed. They were encouraged to use a range of reading strategies whilst doing so and to record them on the checklist. After reading, they showed comprehension by going through the general meaning of the text as a whole class rather than answering the comprehension questions (to keep within the 20–30 minutes allocated). They also reflected on the strategies used and how effective they had been. They were asked to set goals for themselves to achieve in their strategy use and gave each other feedback on their progress towards them. However, unlike in Macaro & Erler (2008), learners did not receive individualized written feedback on strategy use, because participating teachers felt that this would not be possible given time constraints and workload. Nevertheless, the reflection tasks were designed to underline the link between strategy use and understanding the text, that is, to direct learners' attributions for task outcomes towards their strategy use.

Phonics. The phonics intervention again began with some awareness-raising tasks, designed to help students notice and reflect on some of the key differences between French and English SSCs. Subsequently, a set of specific SSCs was introduced each week, using short 'exemplar' words. Learners were also asked to reflect on any other words they knew containing these 'target' SSCs and then to practise reading aloud some sentences containing them.

Finally, every few lessons, learners worked with the same challenging texts as those used in the strategies intervention. In addition to facilitating the deployment of certain strategies, each text had been designed to exemplify particular SSCs. Teachers did not, however, encourage or support learners in the use of strategies or 'sounding out' to comprehend the texts. Instead, learners simply answered the comprehension questions, with brief, whole-class feedback from the teacher.

Texts. The texts group received no explicit instruction in either strategies or phonics. They did, however, work with the same texts as those used in the other two groups. They answered the comprehension questions used with the phonics group, along with the more open-ended questions, and watched the online video clips in class. Again, brief, whole-class teacher feedback was given.

Analysis

We assessed normality and homogeneity of variance assumptions through histograms and normality tests for the variables of interest, for the sample as a whole and for each intervention group, at both time points. Shapiro–Wilks

TABLE 2
Pearson's Correlations Between All Variables

	Self-Efficacy	Reading	LK	TERRS	GAP
Self-efficacy	1	.431**	.266**	.493**	.237**
Reading		1	.369**	.226**	.530**
LK			1	.121*	.268**
TERRS				1	.142**
GAP					1

Note. LK = linguistic knowledge; TERRS = text engagement regulatory reading strategies; GAP = general academic performance.
*$p < .01$. **$p < .001$ (two tailed).

values were significant ($p < .05$) for many variables at both time points, but histograms suggested that deviations from normality were not severe, with values for skewness and kurtosis below 2 (George & Mallery, 2010). Following Field (2013), we therefore felt justified in using parametric statistics (linear mixed effects, ANOVA, and hierarchical cluster analysis; see Results section for full details). In the reporting of results, we follow Plonsky & Oswald's (2014) rules of thumb in interpreting Cohen's *d* effect sizes: between-groups contrasts, small = .4, medium = .7, large = 1.0; and *r*: small = .25; medium = .40, and large = .60. For η^2, we were guided by Plonsky and Ghanbar's (2018) recommendation of .20 and below as small, .50 and above as large effect sizes for second language research. Cohen (1988) suggests that small = .01, medium = .06, large = .14.

RESULTS

Relationship Between L2 Reading Self-Efficacy and L2 Text Engagement Regulatory Reading Strategies

A Pearson correlational analysis was first run between the T1 measures for L2 reading self-efficacy and TERRS for the whole sample, showing a statistically significant relationship of medium strength, $r = .493$, $p < .001$ (two tailed). We also expected that reading self-efficacy would be related to our other variables of interest—as indeed was the case, albeit to a lesser extent than was true for TERRS (see Table 2).

Self-Efficacy for Reading at Time 1. We then examined how far TERRS predicted self-efficacy by running tests of linear mixed effects in SPSS (version 25), which controlled for school-level effects. An initial model assessed how much of the variance in T1 self-efficacy lay between the schools in the sample. This indicated that there was significant variance in intercepts between schools, $\text{var}(u_{0j}) = .80$, Z = 2.20, $p = .014$ (one tailed), meaning that a small but significant portion of the variance (7.47%) in T1 scores was accounted for at the school level. Therefore, a mixed model was fit with school as the upper level and learners at the lower level and with self-efficacy as the outcome variable.

TERRS was next entered as a fixed effect, along with T1 scores for reading comprehension, LK, and GAP (all in the form of z scores), as the initial correlation analyses suggested that these were also possible predictors of reading self-efficacy. The addition of these variables resulted in a significant improvement in overall model fit, $\Delta\ \chi^2(1) = 201.89$, $p < .001$. However, the variance in intercepts between schools was no longer significant. As GAP did not emerge as a significant predictor of reading self-efficacy, $F(1,223.76) = .002$, $p = .97$, it was not retained in the final model. By contrast, TERRS, $F(1,438.80) = 116.51$, $p < .001$ and then reading comprehension, $F(1,139.87) = 55.17$, $p < .001$, emerged as the strongest predictors of reading self-efficacy, followed by LK, $F(1,484.05) = 7.59$, $p = .006$ (see Table 3).

Self-Efficacy for Reading at Time 2. Having established that TERRS significantly predicted reading self-efficacy at T1, as did LK and reading comprehension, we then explored whether this was also the case at T2. As before, a mixed model was fit with school as the upper level and with learners at the lower level. T2 reading self-efficacy scores showed significant variance in intercepts between schools, $\text{var}(u_{0j}) = 1.87$, Z = 2.66, $p = .004$ (one tailed), explaining more variance than at T1 (14.22%). We then added T1 reading self-efficacy, along with the T2 values for those variables that had been entered into the final T1 model, as fixed factors: TERRS, reading comprehension, and LK, significantly improving the overall model fit, $\Delta\ \chi^2(1) = 882.76$, $p < .001$. The variance in

TABLE 3
Model Summary, Time 1 Self-Efficacy for Reading

	Null Model				Final Model					
	σ^2	SE	95% CI	−2LL	b	σ^2	SE	95% CI	−2LL	$\Delta\ \chi^2$ (p)
Fixed effects										
TERRS					1.32^{***}		.12	[1.08, 1.56]		
Reading					$.97^{***}$		.13	[.71, 1.22]		
LK					$.36^{**}$		.13	[.10, .62]		
Random effects										
School	$.80^{*}$	.36	[.33, 1.95]			.07	.14	[.002, 3.35]		
Model fit				2,539.30					2,337.41	201.89^{***}

Note. TERRS = text engagement regulatory reading strategies; LK = linguistic knowledge.
$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .001.$

intercepts between schools was no longer significant once these variables were entered. T1 self-efficacy scores strongly and significantly predicted T2 self-efficacy scores, $F(1{,}353.85) = 56.07$, $p < .001$. As at T1, T2 TERRS was a significant predictor of T2 self-efficacy, $F(1{,}352.79) = 45.27$, $p < .001$, as was T2 reading comprehension, $F(1{,}260.04) = 51.85$, $p < .001$. However, there was no significant effect for T2 LK, $F(1{,}302.003) = .65$, $p = .42$. This time, T2 reading comprehension was the strongest predictor of reading self-efficacy (see Table 4).

Learner Profiles

Having established that reading self-efficacy at T1 was significantly predicted by T1 TERRS, but also by reading comprehension and LK, we then examined whether learners could be categorised into profile types, based on those predictor variables, who would then differ significantly from one another on reading self-efficacy. We used hierarchical cluster analysis, a multivariate exploratory and descriptive procedure which classifies variables or cases into small clusters by observing dissimilarities or distance between the variables (Staples & Biber, 2015). Participants were clustered using Ward's method to reduce potential error from an increase in overall sum of the squared within-cluster distances. T1 TERRS, reading comprehension, and LK scores were used to determine cluster membership (because these were the significant predictors of reading self-efficacy at T1). Reading self-efficacy was then used as the criterion variable. Inspection of the agglomeration schedule indicated a three-cluster solution, whose validity was verified through an ANOVA followed by Bonferroni post hoc tests on both the cluster variables and the criterion variable of reading self-efficacy.

The analysis (see Table 5 and Figure 4) indicated that the three clusters differed significantly from each other on the selected variables, with near-large effect sizes, and on reading self-efficacy (small effect size).

These three profiles can be described as follows.

Cluster 1: The Engaged High Achievers. This cluster had the highest scores on all clustering variables except TERRS (which was still high), which was then accompanied by the highest levels of reading self-efficacy. Only 16% of learners in this cluster had lower-than-average GAP scores.[4]

Cluster 2: The Maximisers. These learners combined the lowest LK scores with the highest

TABLE 4
Model Summary, Time 2 Self-Efficacy for Reading

	Null Model				Final Model					
	σ^2	SE	95% CI	–2LL	*b*	σ^2	SE	95% CI	–2LL	$\Delta\ \chi^2$ (p)
Fixed effects										
T1 self-efficacy					1.15***		.15	[.85, 1.45]		
T2 TERRS					.95***		.14	[.67, 1.22]		
T2 reading					1.24***		.17	[.90, 1.57]		
T2 LK					.12		.15	[–.17, .40]		
Random effects										
School	1.87*	.70	[.89, 3.90]			.0005	.10	[.00, 4.22]		
Model fit				2,524.88					1,642.12	882.76***

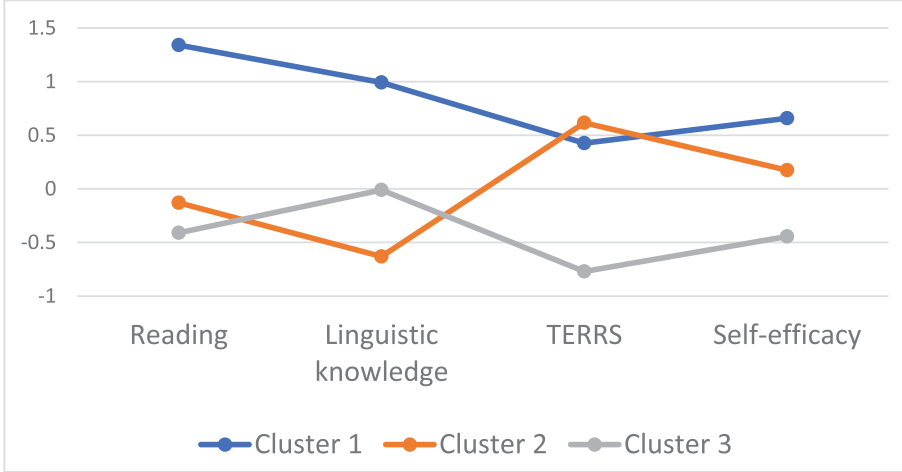
Note. TERRS = text engagement regulatory reading strategies; LK = linguistic knowledge.
*$p < .05$. ***$p < .001$.

TABLE 5
Cluster Profiles

	Cluster 1 ($n = 133$)		Cluster 2 ($n = 179$)		Cluster 3 ($n = 217$)				
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Bonferroni Post Hoc Comparisons	*F*	η squared
T1 TERRS (out of 44)	34.35	4.57	35.45	3.16	27.48	5.10	1, 2 > 3	188.93	.42***
T1 reading (out of 45)	24.79	5.07	11.78	8.05	9.30	6.02	1 > 2 > 3	244.54	.48***
T1 LK (z scores)	1.32	1.03	−.95	.87	−.08	1.22	1 > 3 > 2	172.61	.40***
T1 self-efficacy (out of 24)	17.15	2.09	15.51	3.06	13.42	3.22	1 > 2 > 3	64.89	.21***

Note. SD = standard deviation; T1 = Time 1; TERRS = text engagement regulatory reading strategies; LK = linguistic knowledge.
***$p < .001$.

FIGURE 4
Clusters and Key Variables
[Color figure can be viewed at wileyonlinelibrary.com]



Note. TERRS = text engagement regulatory reading strategies.

reported use of TERRS. This combination was accompanied by low reading comprehension scores; these were, however, significantly higher than those of Cluster 3. Furthermore, their reading self-efficacy scores were higher than might be expected given their levels of attainment. GAP scores in this cluster were mixed, with 46% of learners below average.

Cluster 3: The Withdrawers. This cluster was characterised by the lowest scores on reading self-efficacy, TERRS, and reading comprehension. LK was the second lowest among the three clusters. The cluster had generally lower levels of GAP (below average for 55% of the cluster). Their profile suggested learners with little sense of engagement in reading.

Impact of the Interventions on Different Learner Profile Clusters

The cluster analysis revealed learner groups commencing secondary school language study with distinct profiles, who might then respond in dissimilar ways to different types of reading instruction. We therefore explored the impact of each intervention on T2 reading self-efficacy for each cluster, firstly calculating descriptive statistics at T1 and T2 (see Figure 5). Noticeable is the larger T1–T2 increase for Cluster 3 (*Withdrawers*) in the strategies group (i.e., those who received explicit strategy instruction) compared with the other two intervention groups. Within Cluster 1 (*Engaged High Achievers*), gains in self-efficacy appear more even across the three groups than in other clusters, albeit with a slightly smaller gain for the strategies group compared with the phonics and texts groups.

A two-way ANOVA was conducted on gain scores (used rather than ANCOVA because clusters differed significantly on T1 reading self-efficacy; Maris, 1998), to explore whether they differed by intervention group and cluster membership (with intervention group and cluster as fixed factors). Linear mixed-effects analysis was not used, as school did not contribute significant variance in reading self-efficacy gain scores. There was a statistically significant main effect for intervention group, $F(2,38.69) = 3.76$, $p = .024$, $\eta_p^2 = .017$, but not for cluster, $F(2,5.05) = .49$, $p = .61$, $\eta_p^2 = .002$. The interaction effect between intervention group and cluster membership was, however, statistically significant, $F(4,25.79) = 2.51$, $p = .041$, $\eta_p^2 = .023$, indicating that the effect of the intervention differed across clusters. Levene's test (one of the assumptions for ANOVA) was significant ($p = .032$), meaning some caution is needed in interpreting these results.

Inspection of the profile plot (see Figure 6) suggested that the effect of the strategies intervention was most marked for Cluster 3 and, to a slightly lesser extent, Cluster 2. Tests of simple effects on each cluster in turn with intervention group as the fixed factor suggested that there was no significant effect of group for Cluster 1 and a marginally nonsignificant main effect of group in Cluster 2, $F(2,30.82) = 2.72$, $p = .069$,

FIGURE 5
Time 1–Time 2 Reading Self-Efficacy by Cluster and Intervention Group [Color figure can be viewed at wileyonlinelibrary.com]

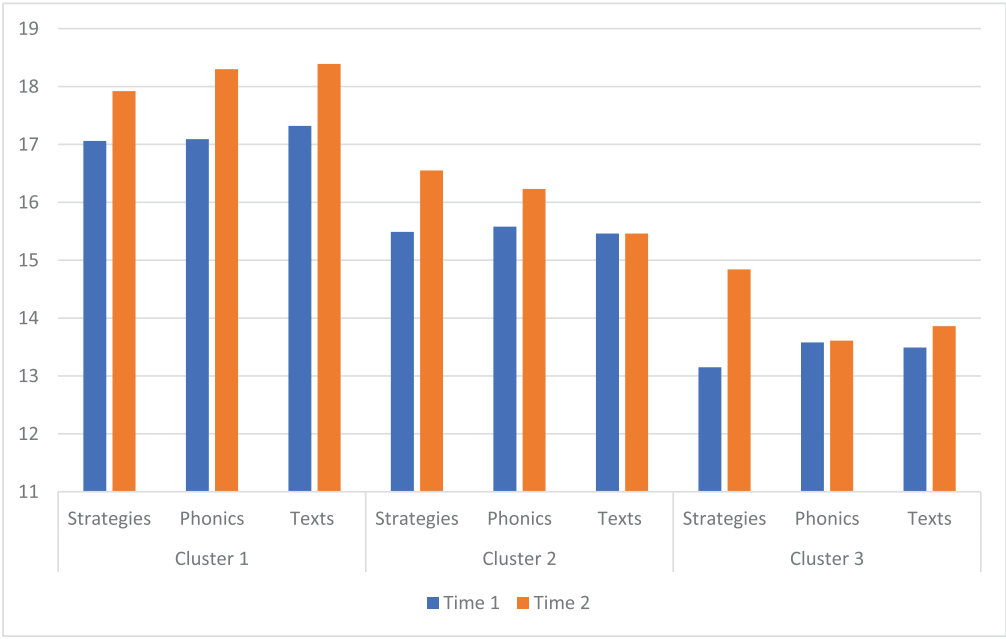
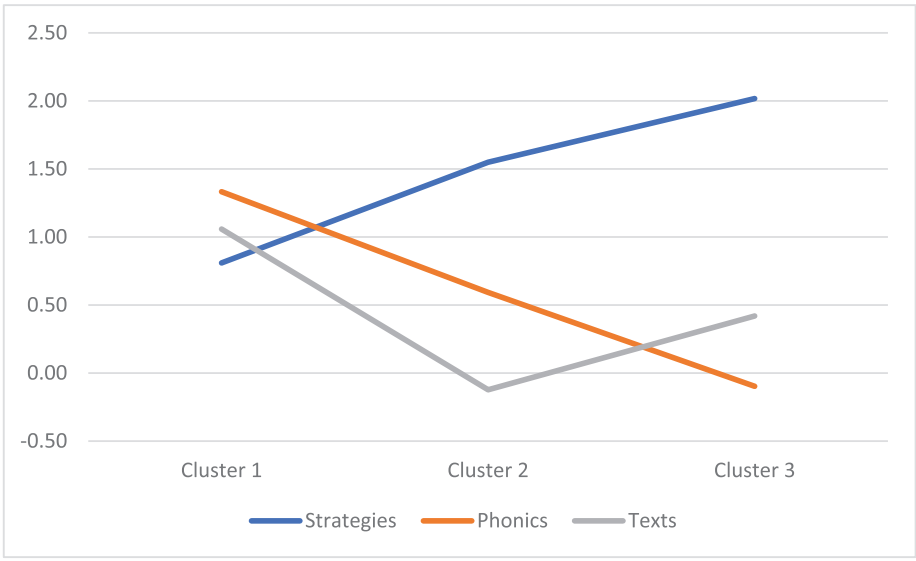


FIGURE 6
Estimated Marginal Means, Self-Efficacy Gains by Cluster and Intervention Group [Color figure can be viewed at wileyonlinelibrary.com]



$\eta_p^2 = .036$. For Cluster 3, there was a significant effect of intervention group, $F(2,74.28) = 6.44$, $p = .002$, $\eta_p^2 = .068$. Post hoc Bonferroni tests showed significant differences between the strategies group and the phonics group, $p = .002$, $d = .64$ (small to medium effect size) and between the strategies group and the texts group, $p = .049$, d = .47 (small effect size).

Finally, we also examined changes in reading comprehension and TERRS use. Although not

FIGURE 7
Time 1–Time 2 Reading Scores by Cluster and Intervention Group [Color figure can be viewed at wileyonlinelibrary.com]

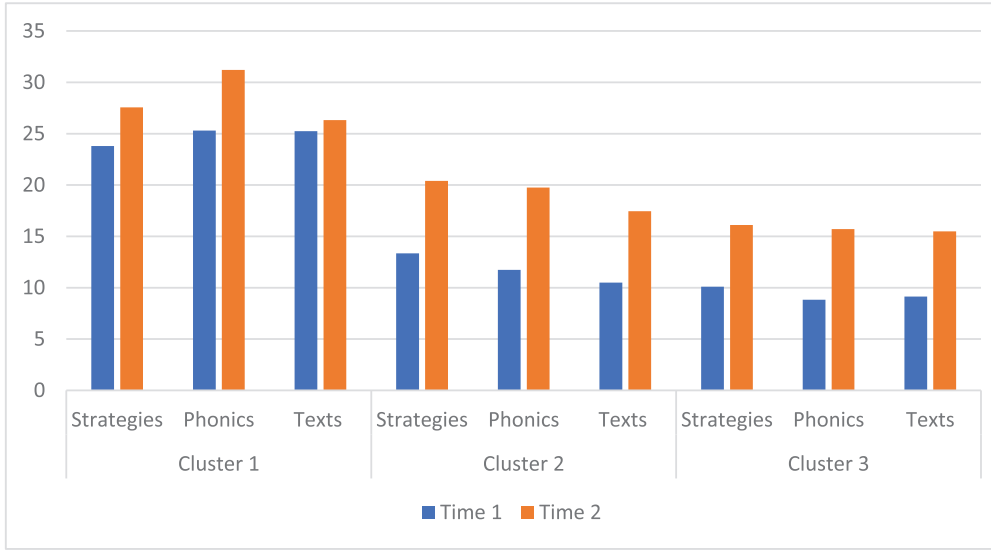
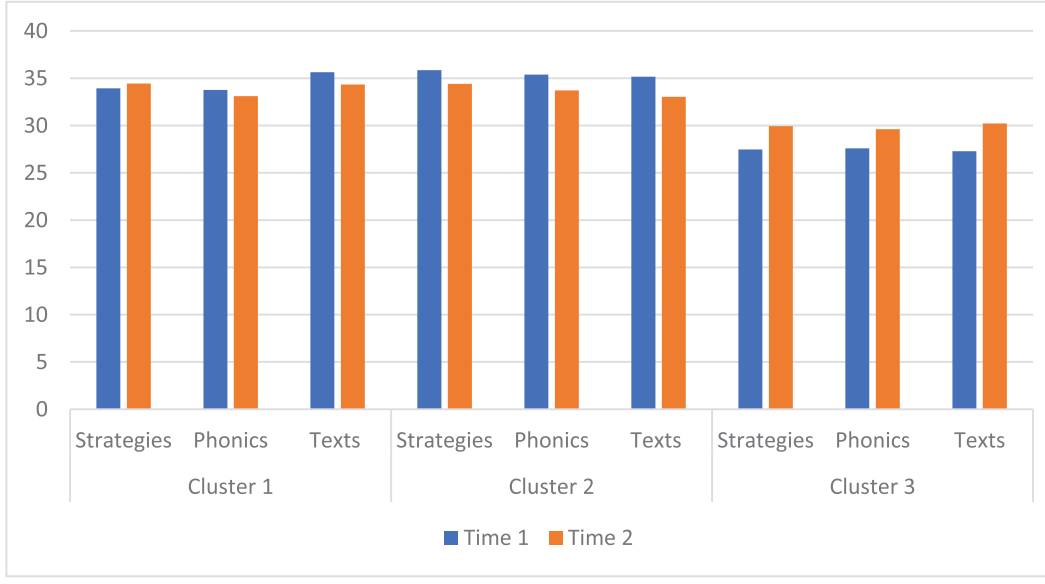


FIGURE 8
Time 1–Time 2 TERRS Scores by Cluster and Intervention Group
[Color figure can be viewed at wileyonlinelibrary.com]



Note. TERRS = text engagement regulatory reading strategies.

directly explored in our research questions, these issues are of relevance for interpreting them. For the 529 learners taken together, the findings reported in Woore et al. (2018) were echoed, as follows: Descriptively, the largest gains for reading were made by the phonics group; for TERRS, it was the strategies group. We also examined whether T1–T2 changes differed by intervention group and cluster membership. From Figures 7 and 8, it can be seen that reading comprehension improved between T1 and T2 across all intervention groups and clusters, with little apparent

variation between them but with generally larger increases for the phonics group. For TERRS, increases appeared only within Cluster 3, across all groups. There was no significant effect of group or cluster on either T2 reading or TERRS scores, controlling for T1 scores.[5]

DISCUSSION

Our findings may be summarised as follows: Among learners of French at the start of secondary school, in a time-poor, limited-L2-input context, a series of analyses showed TERRS to have the strongest relationship with L2 reading self-efficacy among variables that also included L2 reading comprehension, L2 LK, and GAP. Three learner profiles emerged from a cluster analysis, based on their TERRS, L2 LK, and reading comprehension scores at T1; these profile types then differed significantly from one another on reading self-efficacy. While two types of learners were, respectively, high and low on all those variables (Cluster 1, *Engaged High Achievers*, and Cluster 3, *Withdrawers*), a third, Cluster 2 (*Maximisers*), had the lowest scores for LK but relatively high levels of self-efficacy for reading and the highest level of TERRS use. Their reading comprehension scores at T1 were also significantly higher than those of Cluster 3. Furthermore, the impact of the three reading interventions differed by learner cluster. The greatest gains in reading self-efficacy were made by those learners who received reading strategy instruction and who were in Cluster 3, that is, those with lower L2 attainment and GAP. Thus, the study reinforces empirically the place of learner strategies within models of self-regulation such as that of Zimmerman (2013): Strategies, developed through instruction that encourages learners to evaluate their effectiveness and to view them as tools to achieve better outcomes, are an important predictor of self-efficacy, as they may increase learner agency and sense of control. That holds true even when other factors that might influence self-efficacy, such as reading proficiency and LK, are taken into account.

These findings, as well as reinforcing the link between TERRS and self-efficacy development, offer support to those of Magogwe & Oliver (2007), suggesting that strategy development can act as a protective factor for lower proficiency learners who might otherwise withdraw from language learning through lack of confidence, giving them a tool by which they can regulate their performance. This might also be especially relevant for learners whose reading skills in their L1 are not strong. Arguments against strategy-based instruction (such as those from Swan, 2008) claim that when learners reach a certain level of L2 linguistic proficiency, they can access their existing L1 reading strategies, making L2 strategy instruction unnecessary. This argument may, however, have little relevance for learners who have underdeveloped L1 reading strategies in the first place.

Thus, contrary to what might be expected from previous research (Ardasheva et al., 2017), our findings suggest that it is possible to help relatively young learners, with relatively low levels of LK and attainment, to become more confident in reading in another language. Given that learning to read in an L2 requires not only progress in low-level processes such as decoding but also "the ability to engage in reading, to expend effort, to persist in reading without distraction" (Grabe & Jiang, 2013, p. 4), this is not a trivial outcome. Furthermore, the impact of strategy instruction may be more indirect, promoting effort and persistence in the first instance (Teng & Zhang, 2018), with an effect on reading comprehension only becoming apparent over time.

Although descriptively, the strategies group as a whole—and Cluster 3 in particular—increased its TERRS use the most, there was no statistically significant effect of intervention group or cluster on T1–T2 changes. At first glance, this seems surprising given the impact of the strategy intervention on reading self-efficacy and the relationship between TERRS and self-efficacy that we have established. We would argue, however, that by encouraging the strategies group to make the connection between their employment of TERRS and reading outcomes, the intervention did more than teach them strategies—it also helped them to attribute learning to "correctable causes" (Zimmerman, 2000, p. 23), hence fostering a sense of personal agency as an essential part of self-efficacy and of self-regulation more broadly. Our study did not, of course, directly test whether learners' attributions changed as a result of the intervention, and so this explanation remains speculative albeit plausible.

Nevertheless, the increases in self-efficacy in the strategies group were arguably smaller than might have been expected, and smaller than witnessed in Macaro & Erler (2008). Furthermore, the strategies group did not make any greater progress in reading comprehension than the other groups, and descriptively its gains were smaller than those of the phonics group (Woore et al., 2018). In addition, at T2, reading scores themselves were the strongest predictor of reading self-efficacy. It is possible that not making

enough progress in reading—potentially because of the lack of focus on phonological decoding in their instruction—hindered the development of greater self-efficacy within the strategies group. Teachers also reported that some learners found aspects of the strategy intervention difficult, and that they struggled with not having the more familiar type of simple comprehension questions that the other two groups completed. Such questions might have given them what Macaro (2019) called "solid and external evidence that they are indeed making progress" (p. 73). Nevertheless, even though the phonics group made greater gains in reading comprehension at a descriptive level, they showed more modest improvement in self-efficacy, especially among the lower-attaining Cluster 3, for whom phonics alone may have provided insufficient tools to gain access to the challenging texts they encountered. Lower-attaining learners in the phonics group may also have been less able to benefit from the instruction in terms of reading self-efficacy because they were not taught explicitly how to use 'sounding out' as a strategy for comprehension (as they were in Macaro & Erler, 2008). Higher attainers in Cluster 1 perhaps understood how to do this without the need for explicit instruction (thus explaining their greater self-efficacy gains). These findings suggest a need for the combined strategies and phonics approach of Macaro & Erler (2008), but the relative emphasis placed on each type of instruction may need to vary to the meet the needs of learners of different profiles.

LIMITATIONS AND FUTURE STUDIES

Our study is limited by the absence of an analysis of delayed posttest data for our sample, which would have allowed us to explore how durable self-efficacy gains might have been. We also acknowledge that a questionnaire for eliciting data on strategy use offers no insights into actual deployment of strategies on specific tasks. Neither did our methods allow us to explore the important question of whether learners' causal attributions changed as a result of the strategy intervention, which would have allowed us to interpret the relationship between TERRS and reading self-efficacy more fully. These are all areas that would be usefully taken into account in future studies.

CONCLUSION AND IMPLICATIONS

The current study established that the use of TERRS was the most important predictor of reading self-efficacy for beginner learners of French at the start of secondary school, even when reading proficiency and LK were also considered. Three clusters of learners with distinctive profiles emerged, in which TERRS was an important distinguishing factor. While both the *Maximiser* and the *Withdrawer* clusters had low levels of L2 LK alongside generally lower general academic attainment, the former had significantly higher levels of reading self-efficacy, which we attribute to their significantly higher levels of TERRS. They also drew benefit for self-efficacy from our reading strategy intervention, albeit at a marginally nonsignificant level and less so than the *Withdrawers*. The greater effect on the latter may perhaps be because their initially very low levels of TERRS and reading self-efficacy meant they had the most ground to make up.

Taken together, these findings contribute to L2 reading research and theories of self-regulated language learning in three important ways, responding to the call by Rose et al. (2018) for studies "utilizing theory from both self-regulation and language learner strategies" (p. 159) that move the field forward. First, the findings underscore the important cyclical relationship between self-efficacy, strategy use, and learning outcomes, and point to the role of causal attributions in this relationship as a fundamental aspect of self-regulation. By understanding the connection between appropriate use of reading strategies and improved comprehension, learners gain a greater sense of personal agency and heightened self-efficacy, which then leads them to persist longer when reading challenging (but potentially interesting) material and to make effective use of the LK that their instruction is also developing. Second, the relationship between self-efficacy and strategy use is mediated by the different characteristics that learners bring to the classroom, and which instruction needs to take into account. Third, young beginner learners (not only older, more proficient ones) can benefit from reading strategy instruction. While there have been calls to replace language learner strategy research by a focus on self-regulation as a "disposition" rather than investigating "the actual techniques employed" (Tseng, Dörnyei, & Schmitt, 2006, p. 22), such an argument stands in contrast to the wider field, which sees SRL as "skillfully orchestrating processes that are at once covert, behavioral, and environmental" (Zimmerman & Schunk, 2001, p. 304). It also discounts age-related differences in the development of SRL. Young learners such as those in our study, and especially those of lower academic attainment, arguably need reading instruction that does give them some 'actual

techniques' under the guidance of a teacher, albeit within a framework that teaches them how to achieve the necessary orchestration of those techniques and hence develop a self-regulated 'disposition.'

In summary, our findings suggest that a continued focus on language learner strategies, within an overarching self-regulatory framework, is still of value to L2 researchers. It is also of value to practitioners, especially in contexts characterised by heterogeneous learner populations and less-than-favourable conditions for L2 development, where persistence and the ability to deal proactively with setbacks are of prime importance as learners move through different phases of education. Instruction that develops both the lower and higher level processes that underpin successful reading, but that also takes account of the different attributes that learners bring to the classroom, is an area we encourage practitioners to explore in their teaching.

ACKNOWLEDGMENTS

This project was funded by the Nuffield Foundation. The views expressed are those of the authors and not necessarily those of the Foundation. The authors are grateful to all anonymous reviewers and the journal editor, Marta Antón, for constructive feedback on an earlier version of the article. Thanks also go to the teachers and learners involved in the study.

NOTES

[1] Time 3 attrition arose either because the participating teacher moved schools and their replacement did not wish to participate in the project or because classes were regrouped by the school as a result of organisational policies at the start of the new school year.

[2] There was a small amount of missing data for the 529 learners (5% of all values, including 7% and 10% of values for T1 and T2 self-efficacy, respectively). An analysis of the missing values pattern indicated that it was random.

[3] Given the low proficiency of learners, the likelihood of any incidental vocabulary learning from the reading passages was considered very low. We ensured, however, that no words occurred in both the vocabulary and reading comprehension tests (except for *je* 'I,' which occurred in both).

[4] In order to provide further insights into each cluster, we give details of GAP, although this was not used as a clustering variable, because it did not emerge as a significant predictor of T1 reading self-efficacy.

[5] T2 reading scores were analyzed using linear mixed effects, as *school* explained significant and substantial variation in scores. We used ANCOVA for TERRS because no significant effect of school was identified.

Open Research Badges

This article has earned Open Materials badge. Materials are available at https://www.iris-database.org.

REFERENCES

Andon, N., & Wingate, U. (2013). Motivation, authenticity and challenge in German textbooks for Key Stage 3. In J. Gray (Ed.), *Critical perspectives on language teaching materials* (pp. 182–203). Basingstoke, UK: Palgrave Macmillan.

Ardasheva, Y., & Tretter, T. R. (2013). Contributions of individual differences and contextual variables to reading achievement of English language learners: An empirical investigation using hierarchical linear modelling. *TESOL Quarterly*, *47*, 323–351.

Ardasheva, Y., Wang, Z., Adesope, O. O., & Valentine, J. C. (2017). Exploring effectiveness and moderators of language learning strategy instruction on second language and self-regulated learning outcomes. *Review of Educational Research*, *87*, 544–582.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Briggs, J. & Walter, C. (2013). *Read on! Extensive reading and young second language learners' motivation and attitudes*. Oxford: Oxford University Press. Accessed 3 July 2020 at https://ora.ox.ac.uk/objects/uuid:9a696004-725c-4a24-84fb-93c8cec765a1/download_file?safe_filename=Briggs%2526Walter_2016_ReadOn_v2.pdf&file_format=application%2Fpdf&type_of_work=Report

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge Academic.

Courtney, L. (2017). Transition in modern foreign languages: A longitudinal study of motivation for language learning and second language proficiency, *Oxford Review of Education*, *43*, 462–481.

David, A. (2008). Vocabulary breadth in French L2 learners. *The Language Learning Journal*, *36*, 167–180.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*, 231–264.

Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Lawrence Erlbaum.

Erler, L., & Macaro, E. (2011). Decoding ability in French as a foreign language and language

learning motivation. *Modern Language Journal*, *95*, 496–518.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Fung, D., & Macaro, E. (2019). Exploring the relationship between linguistic knowledge and strategy use in listening comprehension. *Language Teaching Research*. https://doi.org/10.1177/1362168819868879

George, D., & Mallery, M. (2010). *SPSS for windows step by step: A simple guide and reference, 17.0 update*. Boston, MA: Pearson.

Grabe, W., & Jiang, X. (2013). Assessing reading. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 1, pp. 185–200). Hoboken, NJ: Wiley.

Grabe, W., & Jiang, X. (2018). First language and second language reading. In J. I. Liontas & M. DelliCarpini (Eds.), *The TESOL encyclopedia of English language teaching*. https://doi.org/10.1002/9781118784235.eelt0476

Graham, S. (2004). Giving up on modern foreign languages? Students' perceptions of learning French. *Modern Language Journal*, *88*, 171–191.

Graham, S., Courtney, L., Tonkyn, A., & Marinis, T. (2016). Motivational trajectories for early language learning across the primary–secondary school transition. *British Educational Research Journal*, *42*, 682–702.

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*, *58*, 747–783.

Gu, Y. (2019). Approaches to learning strategy instruction. In A. U. Chamot & V. Harris (Eds.), *Learning strategy instruction in the language classroom: Issues and implementation* (pp. 22–37). Bristol, UK: Multilingual Matters.

Jeon, E.–Y., & Day, R. R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language*, *28*, 246–265.

Kohler, M. (2017). *Review of languages education policies in Australia*. Report commissioned by the Multicultural Education and Languages Committee. Accessed 3 July 2020 at https://www.education.sa.gov.au/sites/default/files/melc-review-of-languages-education-policies-australia.pdf?acsf_files_redirect

Kormos, J., & Csizér, K. (2014). The interaction of motivation, self-regulatory strategies, and autonomous learning behaviour in different learner groups. *TESOL Quarterly*, *48*, 275–299.

Linnenbrink, E. A., & Pintrich, P. R. (2003) The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly*, *19*, 119–137.

Lonsdale, D., & Le Bras, Y. (2009). *A frequency dictionary of French*. London: Routledge.

Macaro, E. (2019). Language learner strategies and individual differences. In A. U. Chamot & V. Harris (Eds.), *Learning strategy instruction in the language classroom: Issues and implementation* (pp. 68–80). Bristol, UK: Multilingual Matters.

Macaro, E., & Erler, L. (2008). Raising the achievement of young-beginner readers of French through strategy instruction. *Applied Linguistics*, *29*, 90–119.

Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System*, *35*, 338–35.

Maris, E. (1998). Covariance adjustment versus gain scores–revisited. *Psychological Methods*, *3*, 309–332.

Milton, J. (2006). Language lite: Learning French vocabulary in school. *Journal of French Language Studies*, *16*, 187–205.

Murphy, V. (2014). *Second language learning in the early school years: Trends and contexts*. Oxford: Oxford University Press.

National Reading Panel. (2000). *Report of the National Reading Panel–Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*, 422.

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory Into Practice*, *41*, 116–125.

Pawlak, M. (2019). Investigating language learning strategies: Prospects, pitfalls and challenges. *Language Teaching Research*. https://doi.org/10.1177/1362168819876156

Plonsky, L. (2019). Language learning strategy instruction: Recent research and future directions. In A. U. Chamot & V. Harris (Eds.), *Learning strategy instruction in the language classroom: Issues and implementation* (pp. 3–21). Bristol, UK: Multilingual Matters.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912.

Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *Modern Language Journal*, *102*, 713–713.

Raoofi, S., Tan, B. H., & Chan, S. H. (2012). Self-efficacy in second/foreign language learning contexts. *English Language Teaching*, *5*, 60–73.

Rose, H., Briggs, J. G., Boggs J. A., Sergio, L., & Ivanova–Slavianskaia, N. (2018). A systematic review of language learner strategy research in the face of self-regulation. *System*, *72*, 151–163.

Staples, S., & Biber, D. (2015). Cluster analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 243–274). London: Routledge.

Swan, M. (2008). Talking sense about learning strategies. *RELC Journal*, *39*, 262–273.

Taylor, A. M. (2014). L1 glossing and strategy training for improving L2 reading comprehension: A

meta-analysis. *International Journal of Quantitative Research in Education*, *2*, 39–68.

Teng, L. S., & Zhang, L. J. (2018). Effects of motivational regulation strategies on writing performance: A mediation model of self-regulated learning of writing in English as a second/foreign language. *Metacognition and Learning*, *13*, 213–240.

Tinsley, T. (2019). *Language trends 2018: Language teaching in primary and secondary schools in England. Survey report*. London: British Council.

Torgerson, C., & Torgerson, D. (2013). *Randomised controlled trials in education: An introductory handbook*. London: Education Endowment Foundation.

Tseng, W.–T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, *27*, 78–102.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman–Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, *89*, 451–469.

Woore, R., Graham, S., Porter, A., Courtney, L., & Savory, C. (2018). *Foreign language education: Unlocking reading (FLEUR): A study into the teaching of reading to beginner learners of French in secondary school*. Final report to the Nuffield Foundation. Accessed 3 July 2020 at https://ora.ox.ac.uk/objects/uuid:4b0cb239-72f0-49e4-8f32-3672625884f0

Yun, S., Hiver, P., & Al–Hoorie, A. J. (2018). Academic buoyancy: Exploring learners' everyday resilience in the language classroom. *Studies in Second Language Acquisition*, *40*, 805–830.

Zhang, L. J. (2008). Constructivist pedagogy in strategic reading instruction: Exploring pathways to learner development in the English as a second language (ESL) classroom. *Instructional Science: An International Journal of the Learning Sciences*, *36*, 89–116.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). San Diego, CA: Academic Press.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*, 135–147.

Zimmerman, B. J., & Schunk, D. H. (2001). Reflections on theories of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement* (pp. 289–308). Mahwah, NJ: Lawrence Erlbaum Associates.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.