

Assessment of Optimal Pedagogical Factors for Canadian ESL Learners' Reading Literacy Through Artificial Intelligence Algorithms

Ya Xiao¹ & Jie Hu¹

¹ Department of Linguistics and Translation, School of Linguistics and Translation, Zhejiang University, Hangzhou, China

Correspondence: Jie Hu, Department of Linguistics and Translation, School of International Studies, Zhejiang University, Hangzhou, 310058, China. ORCID: <http://orcid.org/0000-0003-2219-2587> E-mail: huj@zju.edu.cn

Received: April 23, 2019 Accepted: May 18, 2019 Online Published: June 12, 2019

doi:10.5539/ijel.v9n4p1

URL: <https://doi.org/10.5539/ijel.v9n4p1>

Abstract

The current study explored the effective pedagogical factors that distinguish high-achieving from low-achieving ESL (English as a second language) primary school learners in reading literacy in Canada. In total, 203 samples (167 high-achieving learners and 36 low-achieving learners from 128 primary schools) in the 4th grade were drawn from the public database of Progress in International Reading Literacy Study (PIRLS) 2016, which is the benchmark for large-scale assessments of reading literacy targeting fourth-grade students. For the first time in the ESL-related research, this study applied an artificial intelligence approach, support vector machine (SVM), to concurrently analyze 41 pedagogical factors associated with reading materials, classroom organization, reading strategies, in-class reading activities and post-reading activities. The overall 41 factors collectively distinguished the high-achieving readers from the low-achieving readers with a high accuracy score (0.793) via SVM. Then, these 41 factors were ranked according to their contribution to the SVM model through SVM-based recursive feature elimination (SVM-RFE). Eventually, an optimal factor set was selected by the SVM-RFE cross validation, which contained 10 effective pedagogical factors centered on reading materials, reading strategies and in-class reading activities for fourth-grade high-achieving ESL learners in reading literacy. Suggestions based on solid data analysis would facilitate infrastructural and pedagogical improvements in ESL reading education.

Keywords: effective pedagogical factors, support vector machine, ESL high-achieving learners, reading literacy, PIRLS 2016

1. Introduction

Pedagogical effectiveness has been enjoying a high profile in reading education because it leads to reading excellence by substantially improving students' literacy achievements in a sustained way (Gallagher, Malloy, & Ryerson, 2016). Compared with reading education for natives, English as a second language (ESL) reading education has a higher demand for teachers' knowledge on pedagogical effectiveness because of ESL students' language disadvantage (Xu, 2015). As one of the most influential large-scale global assessments, Progress in International Reading Literacy Study (PIRLS) regularly assesses the 4th-grade students' reading literacy every five years in the worldwide since 2001. PIRLS provides rich data covering student-, school-, teacher- and parent-levels, and these data have opened up a new area of research on the pedagogical factors of students' reading performance and facilitated the consummation of educational policies (e.g., Hopfenbeck et al., 2017; Lenkeit, Chan, Hopfenbeck, & Baird, 2015; Mullis & Martin, 2015; OECD, 2017). However, the effectiveness of pedagogical factors on learners' reading literacy may vary based on the different language backgrounds of learners; moreover, among current PIRLS studies, ESL-related studies are very few. Moreover, at the 4th-grade level, a high level of reading literacy is a major advantage for future academic achievement while a low level constitutes a difficult obstacle to participating fully in society (Alivernini, 2013). Investigations targeting reading excellence can bring innovative insights to narrow the reading achievement gap. Therefore, the current study aims to identify the effective pedagogical factors that contribute to high achievement among ESL learners based on the PIRLS 2016, which is the latest wave of the test.

1.1 Influence of Pedagogical Factors on ESL Learners' Reading Literacy

Students' language backgrounds play a prominent role in reading literacy. When students are nonnative speakers

of the instruction language, usually an initial gap is formed between the native speakers and the nonnative speakers (Muriel, 2011). Limited language proficiency likely causes processing difficulties of reading materials written in the second language (L2) for L2 learners (Horiba & Fukaya, 2015). Under this circumstance, students' different language backgrounds call for corresponding pedagogies on their reading literacy.

In effect, studies on the pedagogical factors that influence the reading literacy of ESL students have emerged and attracted the attention of educators and researchers. Similar to the research findings on native learners, ESL students' class participation is essential for improving their reading performance because group work is particularly effective in improving ESL students' motivation and engagement in reading in the primary grades. During collaborative activities, such as in-class discussion, students interact with teachers and peers, question the reading materials, hold polite disagreements, enrich each other's arguments, and use textual evidence to convince others (Moses, Ogden, & Kelly, 2015). Unique to ESL learners, oral reading fluency, i.e., the capability to read text with meaningful expression quickly and accurately, is an integral component of ESL primary reading classes. This pedagogy can significantly improve the reading comprehension of ESL learners (Draper & Spaul, 2015). In addition, encouraging students to discuss the understanding of texts orally in class has been proven to achieve better reading results than a systematic phonetic teaching and ICT-assisted reading interventions (Adesope, Lavin, Thompson, & Ungerleider, 2011). Furthermore, building knowledge from the known to the unknown is particularly effective in ESL reading teaching since the activation of students' cultural background knowledge would facilitate meaning construction in reading (Sidhu, Kaur, & Fook, 2018). Therefore, providing reading materials embodied with eastern ideology and teacher-centered pedagogy is especially facilitative in improving the reading literacy of ESL learners from Asia (Beneville & Li, 2018). Additionally, encouraging the development of the first language (L1) reading literacy of ESL learners is also associated with high-level of reading literacy, especially in the case of contrastive analysis with dual-language books.

1.2 Influence of Pedagogical Factors on Reading Literacy in PIRLS

PIRLS is an international assessment of fourth-grade students' reading literacy, which has been launched by the International Association for the Evaluation of Educational Achievement (IEA) every five years since 2001. Reading literacy in PIRLS is defined as the ability to retrieve explicitly stated information, make straightforward inferences, interpret and integrate ideas and information, and evaluate and critique content and textual element (Mullis & Martin, 2015). Students are categorized into 5 levels according to their test scores. Specifically, Level 3 students (scoring between 475 and 550) represent the intermediate benchmark in reading literacy. Levels 4 and 5 students (scoring at and above 550) are high-achieving readers who are above the intermediate international benchmark, while Levels 1 and 2 students (scoring below 475) are low-achieving readers who are below the intermediate benchmark.

The general impact factors contributing to the reading achievement in PIRLS have been explored. Specifically, a diachronic research has been conducted in the Netherlands based on PIRLS 2001, 2006 and 2011 (Netten, Voeten, Droop, & Verhoeven, 2014), and it finds that a decline in reading literacy occurs over time largely because of the gender, socioeconomic status and educational factors of students' early literacy abilities cultivated at home. Additionally, how children are cultivated at home exerts a non-negligible influence on children's reading motivation and self-efficacy, which suggests that parents not only to provide children with rich reading materials but also proper instructions (Cheung et al., 2017). Moreover, the gap of reading literacy between native speakers and immigrants in PIRLS has also attracted research attention. Schnepf (2007) finds that immigrants fare better in reading assessment than natives in Continental European countries caused by school segregation; in contrast, this study also reveals that immigrants receive lower reading scores in English-speaking countries due to inadequate language skills. Lenkeit, Caro and Strand (2015) supplement Schnepf's findings by conducting a study in England and find that the poorer reading performance of immigrants relative to natives is mainly due to their inferior family socioeconomic status.

Research studies on the effectiveness of pedagogical factors that influence reading literacy in PIRLS are far from abundant, and the existing results on this research theme are not consistent. Joseph, Cheung and Raymond (2009) suggest that focusing on using various types of reading materials and teaching different reading strategies may not necessarily improve students' reading literacy and that teachers should arrange courses that can facilitate the identification of low achievers. In contrast to their findings, Cheung, Tse, Lam and Loh (2010) find that instruction in reading strategies is an effective pedagogical factor and teachers' assessment practices and in-service training are helpful for enhancing Hong Kong students' reading literacy in PIRLS 2006. Moreover, the above-mentioned research efforts on PIRLS scarcely distinguish between 4th-grade high achievers and low achievers of reading literacy. However, extreme levels of reading literacy, especially the high-achieving ESL learners, do deserve particular attention since relevant findings would certainly offer innovative insights and lead

to educational excellence for educators and policymakers around the world (Jerrim, 2015). A pioneering study by Alivernini (2013) has shed light on the high-achieving and low-achieving readers in PIRLS with the discovery that the discrepancy in reading performance could be predicted by the interrelationships between student-level factors (e.g., students' reading attitude and self-concept), home-level factors (e.g., home educational resources), school-level factors (e.g., ratio of students from wealthy family) and country-level factors (e.g., primary school teachers' salaries). However, the school-level factors in Alivernini's study (2003) mainly focus on the economic context of the school rather than the pedagogical factors.

1.3 Research Questions

Generally, very few previous studies have examined pedagogical factors influencing extreme levels of reading literacy among primary school students, and none has focused on the ESL learners' reading literacy in PIRLS. Therefore, the current study fills the research gap by focusing on the effective pedagogical factor set distinguishing high- from low-achieving ESL primary school readers. Particularly, students' language background is used as the main independent variable to determine the ESL learners, and it is based on the background information provided by the student questionnaire in PIRLS 2016. Therefore, the ESL cohort is grouped into the high-achieving readers and the low-achieving readers using the PIRLS international reading literacy benchmark. Two research questions are posed in this study: (1) Can fourth-grade ESL learners with high-achieving reading literacy be distinguished from those with low-achieving reading literacy by pedagogical factors? (2) If so, which pedagogical factors compose an optimal factor set with collective effect on high-achieving ESL learners?

2. Materials and Methods

2.1 Sample

The sample was drawn from the PIRLS 2016 dataset (URL: <https://timssandpirls.bc.edu/pirls2016/international-database>), which is the latest PIRLS dataset released in December 2017. Amongst the 50 participating countries and 11 benchmarking regions, 5 majority English-speaking countries (United States, United Kingdom, Canada, Australia and New Zealand) (Nossal, 2012) were all included. It was found that Canada is the only country with valid language background information accessible to the public; therefore, Canada (without Quebec) was taken as the sampling ESL country. According to students' responses of language background in the PIRLS 2016 questionnaire, a total of 203 ESL learners (167 high-achievers and 36 low-achievers) who were not born in Canada (question item of ASBH03A in the PIRLS questionnaire) and whose main language before attending school was not English (question item of ASBH04A in the PIRLS questionnaire) were eventually selected as the samples.

2.2 Variables

In this study, the dependent variables were the reading scores of the two cohorts of students, namely, the high-achieving learners at Levels 4 and 5 and the low-achieving learners at Levels 1 and 2. The independent variables were the coded responses from the teacher questionnaire in PIRLS 2016. Five categories of predictors related to ESL reading teaching were recognized from the PIRLS and included in this study: the choice of reading materials, the organization of the class, the instruction of reading skills or strategies, and the design of in-class reading activities and post-reading activities. The controlling variables included the students' grade, the birth country and the spoken language before entering school. In addition, class size was included as a covariate (see Appendix A).

2.3 Support Vector Machine (SVM) Modeling

2.3.1 Reason for Selecting SVM

Belonged to artificial intelligence, the machine learning algorithm excels in binary classification with complex variables of multiple dimensions, which is an excellent technique in data mining. There are two forms of machine learning techniques: unsupervised (known as descriptive) and supervised (known as predictive). Supervised learning trains a model with labeled input and output data so that it can predict future outputs, whereas unsupervised learning finds underlying meaningful patterns in input data only (Jordan & Mitchell, 2015). In the statistical classification, classification and regression trees (CART) and SVM are two symbolic supervised learning methods of data classification and pattern recognition; however, data analysis in CART is a logic-based analysis, whereas SVM is a rule-based data analysis. CART has been successfully used to analyze PIRLS datasets (Alivernini, 2013), and SVM has recently been effectively utilized in analyzing mathematics literacy in another international benchmark of Program for International Students Assessment (Gorostiaga & Rojo-Álvarez, 2015). CART offers the advantage of providing a straightforward illustration of the correlation

among variables by automatically detecting the complex interaction effects among them (e.g., Alivernini 2013). However, CART has a high potential for over-fitting, or producing a result corresponding too closely or even exactly to a particular dataset, and it may fail to predict additional datasets reliably; moreover, the decision tree of CART is easily affected by small changes to data due to its own hierarchical splitting structure; the classification performance of CART is less consistent than SVM (Shao & Lunetta, 2012). Therefore, SVM was applied in this study because it substantially outperforms CART, which can provide higher overall classification accuracy and less variability due to its robust outlier deletion capability.

2.3.2 Construction of SVM model

SVM is an advanced machine learning algorithm that is superior in classification by constructing a theoretical hyperplane to best separate two groups of multidimensional descriptors, i.e., descriptors with more than one feature (Cortes & Vapnik, 1995). Supposing there is a descriptor, then the theoretical hyperplane is formed by discovering another descriptor v and a parameter that can minimize $|v|^2$ and satisfy the following equations.

$$v \cdot a + q \geq +1 \text{ for positive data} \tag{1}$$

$$v \cdot a + q \leq -1 \text{ for negative data} \tag{2}$$

More intuitively, to divide the descriptors more accurately, the overall distance between all vectors and the hyperplane should be the largest. Namely, the optimal hyperplane should be right in the middle of the vectors closest to this hyperplane. These vectors are the so-called support vectors, and the data are dichotomized using maximum margin formed by support vectors as shown in Figure 1.

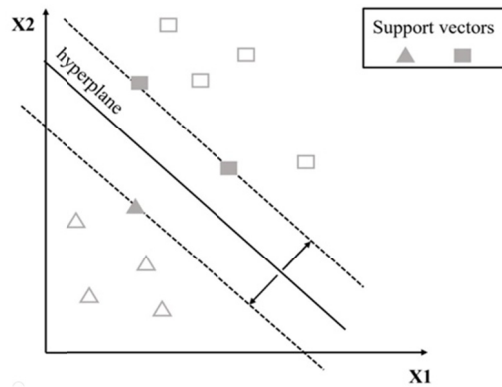


Figure 1. Visualized representation of the support vector machine in two-dimensional space

This optimal hyperplane is found with the kernel function embodied in SVM. When the samples are linearly non-separable, the Mercer kernel for the linear kernel function can be used to map the input samples to a high-order feature space (Chen, Zhang, Wei, & Hu, 2019). Figure 2 visualizes the mapping process realized by the kernel function embodied in SVM.

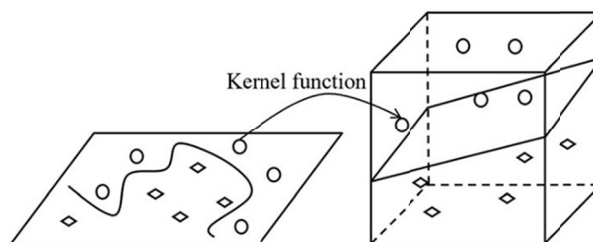


Figure 2. Graphical illustration of the mapping process by the Kernel Function in SVM

To avoid data overfitting, a five-fold cross-validation was implemented in this study. When dealing with multidimensional variables, machine learning algorithms construct a statistical model of sample data, which is called “training data”, to explore the systematical relationships between variables, and to predict a consistent pattern (Larose & Larose 2015). Therefore, the whole dataset was divided into five parts and five SVM models

were implemented. In each fold, each part was predicted one time using the rest four training parts. The formulas for producing the scores of the accuracy, sensitivity and specificity are provided as follows:

$$Accuracy = (TN+TP)/(TP+TN+FP+FN) \quad (3)$$

where:

TN (true negative) stands for the correctly predicted bad performers.

TP (true positive) stands for the correctly predicted good performers.

FP (false positive) stands for the wrongly predicted good performers.

FN (false negative) stands for the wrongly predicted bad performers.

The measurement of accuracy is assessed as the main indicator of the model performance.

SVM has gained wide popularity in the research fields of medical science, computational biology, finance and geoscience. However, the application of machine learning approaches in the humanities is a rare occurrence, which is most likely due to technical difficulty (Jin, Li, Chen, Li, & Hu, 2015), where qualitative analysis is commonly applied to subjects such as corpus linguistics (Chen, Yan, & Hu, 2019; Xiao, Li, & Hu, 2019). Wei, Yang, Chen and Hu (2018) has explored changes in students' learning styles using SVM. Enlightened by their efforts, the current study resorts to LIBSVM (URL: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>), a software package that facilitates the access to SVM by providing many default parameters that prevent certain errors (Chang & Lin, 2011).

Proposed by Guyon, Weston, Barnhill and Vapnik (2002), SVM-recursive feature elimination (SVM-RFE) is an SVM-based feature selection algorithm. RFE is based on the conception to repeatedly construct a new model by removing the weakest feature each time until all features are exhausted. When the final SVM-RFE model is constructed, the contribution of each feature to the formation of the optimal hyperplane can be calculated and then the ranking of features is formed. However, the number of valid features in the final model is unknown in advance. In order to get the certain number of features in the optimal feature set, RFE-cross validation (RFE-CV) was implemented to score different feature collections and select the feature subset with the highest score.

2.4 Data Analysis Procedure

In this study, raw data were preprocessed by first deleting a small number ($N = 30$) of observations with missing values in most columns, integrating the student-level variables and the teacher-level variables via Microsoft Office Access, and then assigning dummy variables 1 and 0 to the high-achieving ESL learners and the low-achieving ESL learners, respectively.

This study conducted a four-step statistical analysis procedure using the publicly free computing program of Python 3.5. First, the whole set of 41 factors were input into the SVM model to examine whether the accuracy score would indicate a distinguishable relationship between the two cohorts (high- and low-achieving learners) caused by pedagogical factors. Second, SVM-RFE was utilized to rank the 41 pedagogical factors according to their contribution to the SVM classification model. Third, RFE-CV was utilized to select the optimal feature set that contributes most to the distinction and visualize the certain number of selected features. Fourth, since teachers' responses were coded from 1 to 4 (1: Every day or almost every day; 2: Once or twice a week; 3: Once or twice a month; and 4: Never or almost never) for the frequency of using these pedagogical practices, the ratio of the recurrence of students' coded answers among high-achieving learners to their recurrence among low-achieving students was computed to indicate the positive or negative influences exerted by pedagogical factors. If the ratio was greater than 1, then this pedagogical factor positively influences high-achieving learners' reading literacy.

3. Results

The overall SVM model achieved an accuracy rate of 0.793, which indicated a strong level of classification between the high- and low-achieving ESL readers. Research question 1 confirmed that pedagogical factors can distinguish the fourth-grade ESL learners with high- and low-achieving reading literacy, which laid the foundation for answering the second research question.

SVM-RFE produced a ranking list of the factors that can distinguish the high- from low-achieving ESL students, according to the factors' different contribution to the distinction. As shown in Figure 3, only the first 17 features showed a stable ranking of discrimination in both of the two ranking patterns produced by SVM-RFE, and the two ranking patterns tended to fluctuate after the 17th feature.

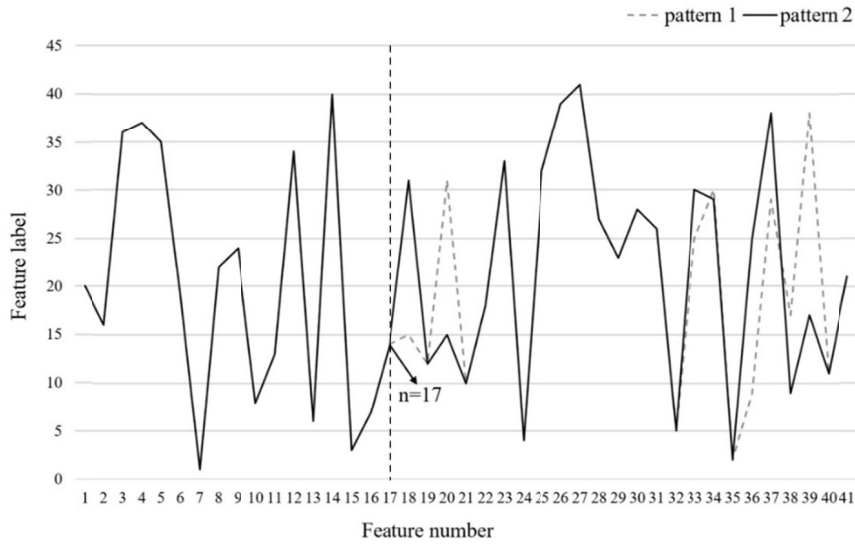


Figure 3. Seventeen features selected from two fluctuate ranking patterns produced by SVM-RFE

To reach the optimal feature set, these 17 features were added to the model in turn to calculate their respective accuracy score by RFE-CV visualizer plot, which is measured by the cross-validated accuracy score. As shown in Figure 4, the curve reached the highest point for the first time when the set of ten informative features was captured; thus, the top-10 features constituted the optimal effective pedagogical factor set with the smallest number of distinguishable features and the highest prediction accuracy (accuracy score = 0.832). Therefore, among the 41 factors, the collective effect of the top-10 features best distinguishes high- from low-achieving ESL readers.

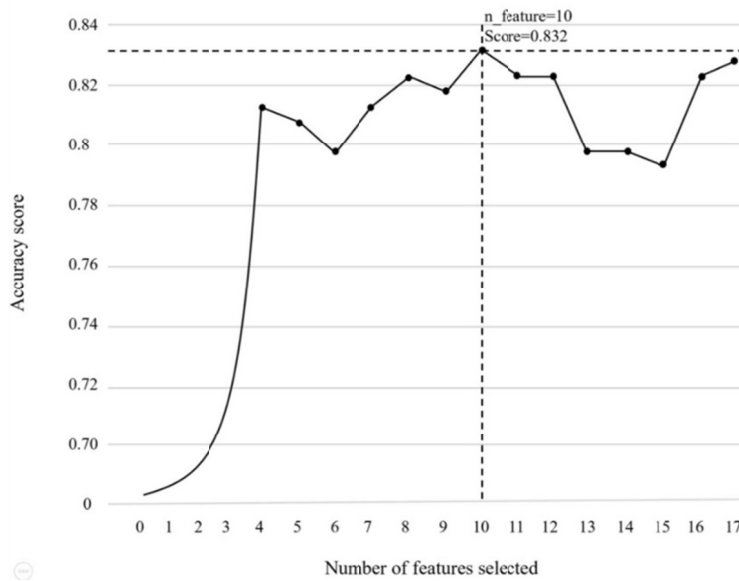


Figure 4. Top-10 features in the optimal feature set selected from 17 features with RFE-CV visualization.

Detailed descriptions of the top-10 significant features and the exact influence of these features are presented in Table 1.

Table 1. Ranking, description and exact influence of the top-10 features in the optimal feature set

Ranking	Top-10 features	Description	Ratio of the recurrence of high-achieving learners' answers to that of low-achieving students' answers		
			Every day (or almost)	Weekly (once or twice)	Monthly (once or twice)
1	ATBR11A	Provide reading materials that match the students' interests	√√ (0.39:0.14=2.8:1)	√ (0.45:0.39=1.15:1)	/ (0.47:0.16=0.34: 1)
2	ATBR10D	Teach students strategies for decoding sounds and words	/ (0.22:0.42=0.52:1)	√ (0.48:0.47=1.01:1)	√ (0.21:0.11=1.89:1)
3	ATBR12H	Describe the style or structure of the text they have read	√√ (0.15:0.06=2.69:1)	√ (0.57:0.42=1.14:1)	/ (0.32:0.44=0.71:1)
4	ATBR12I	Determine the author's perspective or intention	√ (0.12:0.08=1.44:1)	/ (0.46:0.47=0.96:1)	/ (0.34:0.44=0.77:1)
5	ATBR12G	Make generalizations and draw inferences based on what they have read	√ (0.34:0.25=1.34:1)	/ (0.47:0.56=0.85:1)	/ (0.185:0.194=0.95:1)
6	ATBR10G	Teach or model skimming or scanning strategies	√ (0.11:0.08=1.29:1)	/ (0.38:0.50=0.77:1)	√ (0.38:0.36=1.04:1)
7	ATBR01A ^a	Class size ^a	√ ^a (n<20) (0.09:0.06=1.62:1)	√ ^b (20≤n≤30) (0.814:0.805=1.01:1)	√ ^c n>30 (0.10:0.14=0.69:1)
8	ATBR11C	Link new content to students' prior knowledge	√ (0.57:0.56=1.03:1)	/ (0.35:0.36=0.98:1)	/ (0.07:0.08=0.86:1)
9	ATBR11E	Encourage student discussions of texts	√ (0.78:0.75=1.05:1)	√ (0.18:0.11=1.62:1)	/ (0.04:0.14=0.26:1)
10	ATBR09AB	Longer fiction books with chapters	/ (0.50:0.56=0.91:1)	√ (0.31:0.17=1.87:1)	/ (0.14:0.22=0.65:1)

Note. For the covariate, i.e., class size (ATBR01A), code 1 refers to a class size smaller than 20 students (n<20); code 2 refers to a class size of 20 to 30 students (20≤n≤30); and code 3 refers to a class size larger than 30 students (n>30). The ratios larger than 1 are marked by one tick '√', indicating a positive relationship between the pedagogical factors and the high-achieving learners' reading literacy. The ratios larger than 2 are marked by two ticks '√√', indicating a prominent positive relationship.

As demonstrated in Table 1, the components of the optimal feature set of effective pedagogical factors with feature ranking were centered on three of the five categories: reading materials, reading strategies and in-class reading activities. Specifically, regarding the first category of reading material selection, students' interest (ATBR11A) was ranked as the first feature, which indicated that it has the greatest positive influence on predicting high-achieving ESL readers. The type of text also mattered. Compared with informational reading materials, such as nonfiction books and articles, literary reading materials, especially longer fiction books with chapters (ATBR09AB), caused a significant gap between the two cohorts. In terms of the development of reading strategies as the second category, the strategy of decoding sounds (ATBR10D) and the strategy of skimming and scanning (ATBR10G) were found to be strong predictors of high ESL reading performers. In addition, asking students to generalize or infer (ATBR12G), describe the textual style or structure (ATBR12H), and determine the intention of the author (ATBR12I) stood out as the effective methods for improving ESL learners' reading literacy. With regard to the third category of in-class reading activities, considerable discussion of the text (ATBR11E) was indicated as the significant impact factor. Linking teaching content to students' previous knowledge (ATBR11C) was also recommended.

Table 1 further displays how the pedagogical factors of these three categories influence 4th-grade students' reading literacy with frequency. In terms of the frequency of using the top-10 effective pedagogical factors, 7 factors positively influenced high-achieving learners' reading literacy if they are used every day or almost every day, and they include providing reading materials that match the students' interests (ATBR11A with a ratio of 2.8); describing the style or structure of the text they have read (ATBR12H with a ratio of 2.69); determining the author's perspective or intention (ATBR12I, with a ratio of 1.44); making generalizations and draw inferences based on what they have read (ATBR12G, with a ratio of 1.34); teaching skimming or scanning strategies (ATBR10G, with a ratio of 1.29); linking new content to students' prior knowledge (ATBR11C, with a ratio of 1.03); encouraging student discussions of texts (ATBR11E, with a ratio of 1.05). In particular, ATBR11A, ATBR12H and ATBR11E also facilitated high reading literacy when used once or twice a week, with a ratio of 1.15, 1.14, 1.62, respectively; whereas ATBR11A and ATBR12H indicated a less prominent positive influence

compared with their use every day. In addition, teaching strategies for decoding sounds and words (ATBR10D) once or twice weekly (ratio = 1.01) or monthly (ratio = 1.89) predicted high reading literacy, whereas a greater frequency would reverse the positive effect (ratio = 0.52). Longer fiction books with chapters (ATBR09AB) improved high-achieving students' reading literacy only when they were assigned once or twice a week (ratio = 1.87). For the covariate (ATBR01A), a variety of class sizes might lead to different reading performance of ESL readers. A class size smaller than 20 indicated the best reading achievement (ratio = 1.62); a class size from 20 to 30 still contributed to high reading literacy (ratio = 1.01); however, a class size larger than 30 decreased the students' reading achievement (ratio = 0.65).

4. Conclusions and Discussion

In this study, the machine learning method SVM was applied for the first time to identify the optimal feature set of pedagogical factors differentiating 4th-grade ESL learners with high- from those of low-achieving reading literacy. The results identified the top-10 features, which were distributed in three categories (reading materials, reading strategies and in-class reading activities), and class size was identified as the eventual optimal feature set of effective pedagogical factors that can collectively improve the reading literacy of ESL learners. The discovery of this optimal feature set not only highlights the significance of these factors in determining the high-achieving ESL students' reading performance but also corroborates the theoretical premise of pedagogical effectiveness with collective effect. According to the optimal feature set, a conceptual framework with reading pedagogical categories on the reading materials, reading strategies and in-class reading activities were formulated as follows.

4.1 Effective Pedagogical Factors Involved in Reading Materials

In ESL reading teaching, assigning reading materials in line with students' interests was found to be particularly effective in cultivating successful readers in primary school, which is consistent with previous studies that have confirmed the positive influence of instructional materials matched readers' interest areas on L2 reading comprehension (Lien, 2017). The enhancement of reading comprehension may be due to their improvement of L2 proficiency associated with the personalization of the reading materials (Asgari, Ketabi, & Amirian, 2019). Selecting materials that students are interested in fosters intrinsic motivation and encourages extrinsic engagement (Johnson & Blair, 2003). Moreover, fully engaged readers are able to identify with the conceptual context of a text and effortlessly master the meaning of the text (Esther & Lau, 2018). The pedagogy of integrating moral and cultural values of ESL learners' original culture in instructional materials can facilitate reading literacy improvement (Sidhu, Kaur, & Fook, 2018).

In addition, classical research has been conducted to identify the types of reading materials that influence reading comprehension in L2 learning. Carrell (1984a, 1984b, 1985) conducted a series of studies on the influence of different types of rhetorical organization on L2 reading, and in 1984, Carrell found that structured English rhetorical organization could facilitate ESL readers' recalling of the text. Later that year, Carrell discovered that the story structure positively affected ESL reading comprehension by the schema-theoretic approach. Carrell (1985) further studied whether explicit teaching of the structure of reading materials could improve ESL reading and found that the rhetorical structure of expository texts training could greatly increase ESL readers' recalling competence. Another frequent topic in research is comparisons between the memorization and comprehension of narrative and of expository texts. Primary-school readers generally exhibited the same profiles across the two text genres (Josefine et al., 2018). However, in the recall test for memorizing expository texts in L1 and L2, L2 students performed worse than L1 students on the free recall test (Beken & Brysbaert, 2018). Considering literary texts, marginally significant negative correlation has also been observed between non-fiction reading and successful reading comprehension (Topping, Samuels, & Paul, 2008). In this study, longer fiction book reading differentiates high- from low-achieving ESL readers. One possible explanation might be that it is easier to perceive coherence in fiction than non-fiction, which is important because coherence can make texts more inherently interesting and allow readers to connect ideas and make inferences throughout the text (Topping, 2015). In this study, it was also reported that teachers of more high-achieving students' preferred to assign fiction weekly instead of monthly.

4.2 Effective Pedagogical Factors Involved in Reading Strategies

Instruction in reading strategies has always been a focus of L2 teaching. In terms of oral reading strategy, the current study suggests that teaching phonological decoding strategies weekly or monthly predicted better reading performance compared with reading aloud and reading silently. The effectiveness of phonetic decoding has been proved in a longitudinal study targeting students from the first grade to the fourth grade, which found that a lack of phonetic decoding skills was the main obstacle keeping poor readers from improving (Juel, 1988). One possible explanation might be that phonetic decoding facilitates fluency, which has a reciprocal relationship with

reading comprehension (Woore, 2018). However, in existing studies, the pros and cons of reading aloud and silent reading strategies were inconsistent. Sustained silent reading has been suggested to be at least as good as and often better than “regular” instruction focused on skill-building (Krashen, 2005). In particular, silent reading of narrative texts could facilitate retelling. Moreover, reading aloud provides students with a unique opportunity to improve in classroom engagement, and it can increase readers’ vocabulary, reading fluency and speed, and higher-level thinking (Johnston, 2016). In contrast, both oral reading and silent reading of expository texts were shown to play no significant role in comprehension improvement (Schimmel & Ness, 2016).

Skimming and scanning are two significant reading strategies often used by excellent ESL readers, which is consistent with previous studies. Brown (1994) even proposed that skimming and scanning are probably the two most valuable reading strategies both for ESL learners and for native speakers. Fauzi (2018) also found that skimming and scanning have positive influence on reading comprehension through an experiment with five sessions of treatments. The advantage of skimming and scanning strategies still lies in speed reading. Skimming is three to four times faster than normal reading, and people often skim when dealing with lots of materials (Hong, 2013). Scanning is also an effective time-saving strategy for locating specific information that frequently appears in an assignment or an examination (Schlosser, Wendt, Bhavnani, & Nail-Chiwetalu, 2011). Therefore, students who have a good command of these two strategies may achieve high scores in reading tests.

In addition, tasks involving drawing inferences, describing the style or structure of the text, and determining the author’s perspective were found to significantly improve proficient reading development. In effect, the tasks assess students’ metacognitive reading strategies, such as inferring and summarizing, which provides a significant contribution to reading achievement. Inferring includes establishing meaningful connections between individual pieces of literally stated information in the text (i.e., “text-connecting” inferences) and between literally-stated information and readers’ background knowledge (i.e., “knowledge-based” inferences). Successfully drawing inferences can help the readers determine the central ideas of the authors (Hall & Barnes, 2017). Summarizing is constructing a macrostructure by preserving the general meaning and structure of original text, which can also facilitate students’ understanding of information and transferring it to long-term memory (Serpil, 2018).

4.3 Effective Pedagogical Factors Involved in In-Class Reading Activities

The pedagogy of in-class activities has a direct influence on the effectiveness of ESL reading teaching. Linking new content to students’ prior knowledge was predictive of high-achieving ESL readers, which is consistent with previous studies. In-class linking with prior knowledge was found to effectively improve reading comprehension and particularly facilitate memorizing. For instance, adapted from Jean Piaget’ schema theory, Abdelaal and Sase (2014) drew three types of schema, i.e., linguistic schema, formal schema and content schema through a prior-knowledge questionnaire, in which prior knowledge structures are higher-level constructs encompassing commonalities across events. In addition, new information could be integrated into a structure provided by prior knowledge. Educators are advised to activate students’ prior knowledge for procedural knowledge (Courtney, Lauren, Loughlin, & Patricia, 2015; Hu & Wei, 2018) because prior knowledge does not automatically benefit the memory processing of incoming information.

Classroom discussions were also considered an effective pedagogy for improving ESL students’ reading literacy in this study. Students were often undervalued in their interpretation abilities of reading materials such as contemporary picture books (Pantaleo, 2008). Indeed, encouraging students to discuss in class provides them the chance to scaffold each other’s understanding of the text. Through collaborative classroom discussions, students would think critically, clarify their understanding, use comprehension skills and communicate knowledge (Chen & Hu, 2018). In ESL teaching, collaborative discussions serve particularly valuable purposes by allowing students with different backgrounds to participate rich discussions from a transactional perspective. Therefore, educators are supposed to make their expectations of the discussion groups explicit in order to successfully engage each student to classroom discussions. Also, educators can design various types of discussion groups to facilitate meaningful discussions in the primary grades (Moses, Ogden, & Kelly, 2015).

4.4 Effective Pedagogical Factor of Class Size

In this study, the class size bigger than 30 would cause a sharp decrease in students’ reading achievement. This research finding was consistent with previous studies. For instance, Nandrup (2016) conducted a longitudinal research on Danish public school, and found that a class size bigger than 28 would lead to a decrease in students’ reading literacy. Additionally, the increased class size caused decreased reading achievement of ESL learners (Krassel & Heinesen, 2014), which might be that teachers in large classes may not be able to apply a variety range of pedagogies and would instead chose to utilize the teacher-centered pedagogies to organize the large

class (Almulla, 2015), which might account for students' low reading literacy.

Regarding the future research, different pedagogical patterns utilized for L1 and L2 readers should be compared if reliable data for a large population can be collected and the data mining technique applied in this study could be combined with other statistical tools, e.g., regression, to perform more scientific and systematic pedagogical influence analyses in depth.

Acknowledgments

Authors gratefully acknowledge the research project supported by the Zhejiang Federation of Humanities and Social Sciences Circles (Survey Project). Authors would like to express sincere gratitude towards IEA for its generous publication of the data and valuable suggestions on the methods of dealing with the data.

References

- Abdelaal, N. M., & Sase, A. S. (2014). Relationship between prior knowledge and reading comprehension. *Advances in Language & Literary Studies*, 5(6), 125–131. <https://doi.org/10.7575/aiac.all.v.5n.6p.125>
- Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2011). Pedagogical strategies for teaching literacy to ESL immigrant students: A meta-analysis. *British Journal of Educational Psychology*, 81(4), 629–653. <https://doi.org/10.1111/j.2044-8279.2010.02015.x>
- Alivernini, F. (2013). An exploration of the gap between highest and lowest ability readers across 20 countries. *Educational Studies*, 39(4), 399–417. <https://doi.org/10.1080/03055698.2013.767187>
- Almulla, M. A. (2015). An Investigation of Teachers' Perceptions of the Effects of Class Size on Teaching. *International Education Studies*, 8(12), 33–42. <https://doi.org/10.5539/ies.v8n12p33>
- Asgari, M., Ketabi, S., & Amirian, Z. (2019). Interest-Based Language Teaching: Enhancing Students' Interest and Achievement in L2 Reading. *Iranian Journal of Language Teaching Research*, 7(1), 61–75.
- Beken, H. V., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition*, 21(5), 1062–1074. <https://doi.org/10.1017/S1366728917000189>
- Beneville, M. A., & Li, C. (2018). Evidence-based literacy interventions for East/Southeast Asian English language learners: A review of the research and recommendations for practice. *Journal for Multicultural Education*, 12(1), 50–66. <https://doi.org/10.1108/JME-12-2016-0061>
- Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. Upper Saddle River, Prentice Hall Regents.
- Carrell, P. L. (1984a). The effects of rhetorical organization on ESL readers. *TESOL Quarterly*, 18, 441–469. <https://doi.org/10.2307/3586714>
- Carrell, P. L. (1984b). Evidence of a formal schema in second language comprehension. *Language Learning*, 34, 87–112. <https://doi.org/10.1111/j.1467-1770.1984.tb01005.x>
- Carrell, P. L. (1985). Facilitating ESL reading by teaching text structure. *TESOL Quarterly*, 19, 727–752. <https://doi.org/10.2307/3586673>
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27. <https://doi.org/10.1145/1961189.1961199>
- Chen, J., & Hu, J. (2018). Enhancing L2 learners' critical thinking skills through a connectivism-based intelligent learning system. *International Journal of English Linguistics*, 8(6), 12–21. <https://doi.org/10.5539/ijel.v8n6p12>
- Chen, J., Zhang, Y., Wei, Y., & Hu, J. (2019). Discrimination of the contextual features of top performers in scientific literacy using a machine learning approach. *Research in Science Education*. Retrieved from <https://rdcu.be/btN56>. <https://doi.org/10.1007/s11165-019-9835-y>
- Chen, X., Yan, Y., & Hu, J. (2019). A corpus-based study of Hillary Clinton's and Donald Trump's linguistic styles. *International Journal of English Linguistics*, 9(3), 13–22. <https://doi.org/10.5539/ijel.v9n3p13>
- Cheung, W. M., Lam, J. W. I., Au, D. W. H., So, W. W. Y., Huang, Y., & Tsang, H. W. H. (2017). Explaining student and home variance of Chinese reading achievement of the PIRLS 2011 Hong Kong. *Psychology in the Schools*, 54(9), 889–904. <https://doi.org/10.1002/pits.22041>
- Cheung, W. M., Tse, S. K., Lam, J. W. I., & Loh, E. K. Y. (2010). Progress in international reading literacy study 2006 (PIRLS): Pedagogical correlates of fourth - grade students in Hong Kong. *Journal of Research in*

- Reading*, 32(3), 293–308. <https://doi.org/10.1111/j.1467-9817.2009.01395.x>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Courtney, H. S., Lauren, M., Loughlin, S., & Patricia, A. A. (2015). Prior Knowledge Activation in Design and in Practice. *Literacy Research: Theory, Method, and Practice*, 64(1), 478–497. <https://doi.org/10.1177/2381336915617603>
- Draper, K., & Spaul, N. (2015). Examining Oral Reading Fluency among Rural Grade 5 English Second Language (ESL) Learners in South Africa: An Analysis of NEEDU 2013. *South African Journal of Childhood Education*, 5(2), 44–77. <https://doi.org/10.4102/sajce.v5i2.390>
- Esther, S. C. H., & Lau, K. L. (2018). Reading engagement and reading literacy performance: effective policy and practices at home and in school. *Journal of Research in Reading*, 41(4), 657–679. <https://doi.org/10.1111/1467-9817.12246>
- Fauzi, I. (2018). The effectiveness of skimming and scanning strategies in improving comprehension and reading speed rates for the students of English study program. *Register Journal*, 11(1), 75–90. <https://doi.org/10.18326/rgt.v11i1.101-120>
- Gallagher, M. J., Malloy, J., & Ryerson, R. (2016). Achieving excellence: bringing effective literacy pedagogy to scale in Ontario's publicly-funded education system. *Journal of Educational Change*, 17(4), 477–504. <https://doi.org/10.1007/s10833-016-9284-6>
- Gorostiaga, A., & Rojo-Álvarez, J. L. (2015). On the use of conventional and statistical-learning techniques for the analysis of PISA results in Spain. *Neurocomputing*, 171, 625–637. <https://doi.org/10.1016/j.neucom.2015.07.001>
- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 389–422. <https://doi.org/10.1023/A:1012487302797>
- Hall, C., & Barnes, M. A. (2017). Inference instruction to support reading comprehension for elementary students with learning disabilities. *Intervention in School and Clinic*, 52(5), 279–286. <https://doi.org/10.1177/1053451216676799>
- Hong, N. C. (2013). Teaching of skimming at tertiary level: Theoretical and pedagogical issues. *International Journal of Bilingual & Multilingual Teachers of English*, 1(1), 1–7. <https://doi.org/10.12785/ijbmt/010101>
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J. A. (2017). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research*, 3, 333–353. <https://doi.org/10.1080/00313831.2016.1258726>
- Horiba, Y., & Fukaya, K. (2015). Reading and learning from L2 text: Effects of reading goal, topic familiarity, and language proficiency. *Reading in a Foreign Language*, 27(1), 22–46. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1059624.pdf>
- Hu, J., & Wei, Y. (2018). The centrality of creativity: A new perspective on English language teaching. *English Today*, 35(2), 60–62. <https://doi.org/10.1017/S0266078418000299>
- Jerrim, J. (2015). Why do East Asian children perform so well in PISA? An investigation of Western-born children of East Asian descent. *Oxford Review of Education*, 41(3), 310–333. <https://doi.org/10.1080/03054985.2015.1028525>
- Jin, Y., Li, B., Chen, N., Li, X., & Hu, J. (2015). The discrimination of learning styles by bayes-based statistics: An extended study on ILS system. *Control and Intelligent Systems*, 43(2), 68–75. <https://doi.org/10.2316/Journal.201.2015.2.201-2666>
- Johnson, D., & Blair, A. (2003). The importance and use of student self-selected literature to reading engagement in an elementary reading curriculum. *Reading Horizons*, 43(3), 180–202. Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol43/iss3/3
- Johnston, V. (2016). Successful read-alouds in today's classroom. *Kappa Delta Pi Record*, 52(1), 39–42. <https://doi.org/10.1080/00228958.2016.1123051>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>

- Josefine, K., Paul, A. D. B., Anne, H., Marian, H., Arnout, K., & Linda, V. L. (2018). Profiles of young readers: Evidence from thinking aloud while reading narrative and expository texts. *Learning and Individual Differences, 67*, 105–116. <https://doi.org/10.1016/j.lindif.2018.08.001>
- Joseph, W. I., Cheung, W. M., & Raymond, Y. H. (2009). Learning to read. *Chinese Education & Society, 42*(3), 6–32. <http://doi.org/10.2753/CED1061-1932420301>
- Juel, C. (1988). Learning to read and write: a longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437–447. <https://doi.org/10.1037/0022-0663.80.4.437>
- Krashen, S. (2005). Is in-school free reading good for children? Why the national reading panel report is (still) wrong? *Phi Delta Kappan, 86*(6), 444–447. <https://doi.org/10.1177/003172170508600607>
- Krassel, K. F., & Heinesen, E. (2014). Class-size effects in secondary school. *Education Economics, 22*(4), 412–426. <https://doi.org/10.1080/09645292.2014.902428>
- Larose, D. T., & Larose, C. D. (2015). *Data Mining and Predictive Analytics* (2nd ed.). Hoboken: John Wiley & Sons.
- Lenkeit, J., Caro, D. H., & Strand, S. (2015). Tackling the remaining attainment gap between students with and without immigrant background: An investigation into the equivalence of SES constructs. *Educational Research and Evaluation, 21*(1), 60–83. <https://doi.org/10.1080/13803611.2015.1009915>
- Lenkeit, J., Chan, J., Hopfenbeck, T. N., & Baird, J. A. (2015). A review of the representation of PIRLS related research in scientific journals. *Educational Research Review, 16*, 102–115. <https://doi.org/10.1016/j.edurev.2015.10.002>
- Lien, H. Y. (2017). EFL college learners' perceptions of self-selected materials for extensive reading. *The English Teacher, 39*, 194–204. Retrieved from <http://journals.melta.org.my/index.php/tet/article/view/279/176>
- Moses, L., Ogden, M., & Kelly, L. B. (2015). Facilitating meaningful discussion groups in the primary grades. *Reading Teacher, 69*(2), 233–237. <https://doi.org/10.1002/trtr.1392>
- Mullis, I. V. S., & Martin, M. O. (Ed.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Boston, MA: TIMSS & PIRLS International Study Center.
- Muriel, M. (2011). Immigration and student achievement: Evidence from Switzerland. *Economics of Education Review, 30*(1), 16–38. <https://doi.org/10.1016/j.econedurev.2010.06.017>
- Nandrup, A. B. (2016). Do class size effects differ across grades? *Education Economics, 24*(1), 83–95. <https://doi.org/10.1080/09645292.2015.1099616>
- Netten, A., Voeten, M., Droop, M., & Verhoeven, L. (2014). Sociocultural and educational factors for reading literacy decline in the Netherlands in the past decade. *Learning and Individual Differences, 32*, 9–18. <https://doi.org/10.1016/j.lindif.2014.02.002>
- Nossal, K. R. (2012). The Anglosphere: A genealogy of a racialized identity in international relations by Srdjan Vucetic. *International Affairs, 88*(1), 163–165. <https://doi.org/10.1177/002070201206700119>
- OECD. (2017). *PISA 2015 technical report*. Paris, France: OECD Publishing.
- Pantaleo, S. (2008). *Exploring student response to contemporary picturebooks*. Toronto, Canada: University of Toronto Press. <https://doi.org/10.3138/9781442688063>
- Schimmel, N., & Ness, M. (2016). The effects of oral and silent reading on reading comprehension. *Reading Psychology, 38*(4), 1–27. <https://doi.org/10.1080/02702711.2016.1278416>
- Schlosser, R. W., Wendt, O., Bhavnani, S., & Nail-Chiwetalu, B. (2011). Use of information-seeking strategies for developing systematic reviews and engaging in evidence-based practice: The application of traditional and comprehensive pearl growing. a review. *International Journal of Language & Communication Disorders, 41*(5), 567–582. <https://doi.org/10.1080/13682820600742190>
- Schnepf, S. V. (2007). Immigrants' educational disadvantage: An examination across ten countries and three surveys. *Journal of Population Economics, 20*(3), 527–545. <https://doi.org/10.1007/s00148-006-0102-y>
- Serpil, O. (2018). The Effect of Summarization Strategies Teaching on Strategy Usage and Narrative Text Summarization Success. *Universal Journal of Educational Research, 6*(10), 2199–2209. <https://doi.org/10.13189/ujer.2018.061018>
- Shao, Y., & Lunetta, R. S. (2012). Comparison of support vector machine, neural network, and cart algorithms for

- the land-cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing*, 70, 78–87. <https://doi.org/10.1016/j.isprsjprs.2012.04.001>
- Sidhu, G. K., Kaur, S., & Fook, C. Y. (2018). Analysing the integration of moral and cultural values through ELT reading materials in Malaysian ESL classrooms. *Situating Moral and Cultural Values in ELT Materials*, 9, 69–85. https://doi.org/10.1007/978-3-319-63677-1_5
- Topping, K. J. (2015). Fiction and Non-Fiction Reading and Comprehension in Preferred Books. *Reading Psychology*, 36(4), 350–387. <https://doi.org/10.1080/02702711.2013.865692>
- Topping, K. J., Samuels, J., & Paul, T. (2008). Independent reading: The relationship of challenge, non-fiction and gender to achievement. *British Educational Research Journal*, 34(4), 505–524. <https://doi.org/10.1080/01411920701609380>
- Wei, Y., Yang, Q., Chen, J., & Hu, J. (2018). The exploration of a machine learning approach for the assessment of learning styles changes. *Mechatronic Systems and Control*, 46(3), 121–126. <https://doi.org/10.2316/Journal.201.2018.3.201-2979>
- Woore, R. (2018). Learners' pronunciations of familiar and unfamiliar French words: What can they tell us about phonological decoding in an L2? *Language Learning Journal*, 46(4), 456–469. <https://doi.org/10.1080/09571736.2016.1161062>
- Xiao, Y., Li, Y., & Hu, J. (2019). Construction of the Belt and Road Initiative in Chinese and American media: A critical discourse analysis based on self-built corpora. *International Journal of English Linguistics*, 9(3), 68–77. <https://doi.org/10.5539/ijel.v9n3p68>
- Xu, W. (2015). Exploring ESL/EFL teachers' pedagogical content knowledge on reading strategy instruction. *English Language Teaching*, 8(11), 155–175. <https://doi.org/10.5539/elt.v8n11p155>

Appendix A

Pedagogical Variables Included in the PIRLS 2016 Questionnaire

Categories	Label	Description
Variables		
Organizing the class	ATBR08A	I teach reading as a whole-class activity
	ATBR08B	I create same-ability groups
	ATBR08C	I create mixed-ability groups
	ATBR08D	I use individualized instruction for reading
	ATBR08E	Students work independently on an assigned plan or goal
Choosing reading materials	ATBR09AA	Short stories (e.g., fables, fairy tales, action stories, science fiction, detective stories)
	ATBR09AB	Longer fiction books with chapters
	ATBR09AC	Plays
	ATBR09BA	Nonfiction subject area books or textbooks
	ATBR09BB	Longer nonfiction books with chapters
	ATBR09BC	Nonfiction articles that describe and explain about things, people, events, or how things work (e.g., newspaper articles, brochures)
	ATBR11A	Provide reading materials that match the students' interests
In-class reading activities	ATBR11B	Provide materials that are appropriate for the reading levels of individual students
	ATBR11C	Link new content to students' prior knowledge
	ATBR11D	Encourage students to develop their understandings of the text
	ATBR11E	Encourage student discussions of texts
	ATBR11F	Encourage students to challenge the opinion expressed in the text
	ATBR11G	Use multiple perspectives (among students and texts) to enrich understanding
	ATBR11H	Give students time to read books of their own choosing
Developing reading skills or strategies	ATBR11I	Give individualized feedback to each student
	ATBR10A	Read aloud to students
	ATBR10B	Ask students to read aloud
	ATBR10C	Ask students to read silently on their own
	ATBR10D	Teach students strategies for decoding sounds and words
	ATBR10E	Teach students new vocabulary systematically
	ATBR10F	Teach students how to summarize the main ideas
	ATBR10G	Teach or model skimming or scanning strategies
	ATBR12A	Locate information within the text
	ATBR12B	Identify the main ideas of what they have read
	ATBR12C	Explain or support their understanding of what they have read

	ATBR12D	Compare what they have read with experiences they have had
	ATBR12E	Compare what they have read with other things they have read
	ATBR12F	Make predictions about what will happen next in the text they are reading
	ATBR12G	Make generalizations and draw inferences based on what they have read
	ATBR12H	Describe the style or structure of the text they have read
	ATBR12I	Determine the author's perspective or intention
Post-reading activities	ATBR13A	Write something about or in response to what they have read
	ATBR13B	Answer oral questions about or orally summarize what they have read
	ATBR13C	Talk to each other about what they have read
	ATBR13D	Take a written quiz or test about what they have read
Covariates		
	ATBR01A	Number of students in the class
Controlling variables		
	IDGRADE	Grade
	ASBH03A	Whether be born in the test country
	ASBH04A	Spoken language before going to school

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).