Correlation Analysis between Early Reading Amount and Expressive Ability of Young Children Aided by Multiple Information Processing Techniques under the Heuristic Pattern

Fude Duan,1 Jianhua Zhou,2 and Peng Li3

1Department of Chinese Literature, Ganzhou Teachers College, Ganzhou 341000, China
2Institute of Education, Xiamen University, Xiamen, Fujian 361005, China
3Department of Chinese Literature, Ganzhou Teachers College, Ganzhou, Jiangxi 341000, China

Correspondence should be addressed to Jianhua Zhou; zhoujh@stu.xmu.edu.cn

Received 1 July 2022; Revised 9 August 2022; Accepted 20 August 2022; Published 27 September 2022

Academic Editor: Hangjun Che

Copyright © 2022 Fude Duan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The cultivation of reading ability runs through a child’s life. Children are in a critical period of early reading ability formation. Only based on the actual development characteristics of children and adopting effective training methods, can early childhood educators better meet the needs of children’s long-term development. Based on the data of 22 classes in 7 public kindergartens in X City in Mainland of China, this study first compared the effects of two different teaching modes and then analyzed the relationship between early reading amount and expressive ability of children. Independent sample T test, OLS regression, and Lasso regression were used to demonstrate the correlation between early reading amount and expressive ability, and the possible mechanisms of family background are also explored. The results show the following: firstly, heuristic teaching mode has positive teaching effect. The effect quantities of heuristic teaching mode on transmission ability, communication ability, transfer ability, and generalization ability are 0.21, 0.25, 0.31, 0.30, and 0.32, respectively, which belong to medium intensity, indicating that heuristic teaching mode has a good intervention effect on improving students’ expression ability. Secondly, there is a high correlation between early reading and children’s expressive ability. Thirdly, there are two aspects of the influence of early reading amount on children’s expressive ability: explicit influence and implicit influence. Fourthly, the association between early reading and different presentation skills is not the same. To sum up, we should actively encourage the use of heuristic teaching mode, realize the heterogeneity of early reading amount, and promote the improvement of children’s expressive ability through both explicit and implicit paths.

1. Introduction

Expressive ability is one of the common ability young children rely on to communicate with others. It is the ability to express one’s thoughts, feelings, ideas, and intentions clearly and unambiguously in words, characters, graphics, expressions, gestures, etc. The better the ability to express oneself, the better others can understand and appreciate what one wants to express. Previous research has explored the influencing factors of children’s expression ability from multiple angles. For preschool children, expression ability is closely related with their family environment, which is very important to the children’s development in the future. The focus on expressive ability has rich education significance and great research value. Existing research shows that different education methods and educational ideas will have different effects on the development of children’s expression ability. Therefore, for kindergartens and parents, it is worth pondering how to take effective educational measures to promote the complete and all-round development of children’s expression ability. Unfortunately, no researchers have explored children’s expressive ability based on longitudinal follow-up data, and the existing correlation evidence has failed to control the influence of some interfering factors [1].

Based on previous theory and practice, this research will adopt three-phase tracking data, perform in-depth analysis
of the influence factors of children’s expressive ability, and at the same time, based on the perspective of early reading, analyze the key influencing factors and correlation coefficient in mechanism, and provide valuable advice for the practical level of education workers.

In addition, the data of this study are mainly from front-line practical classes. After teaching, teachers use the remaining 10 to 15 minutes of class to measure children’s expressive ability and store the statistic in the data system for data analysis. Students’ family background and demographic data are collected from the information platform of the school and matched based on students’ IDs.

Teaching models have a significant impact on young children’s representational skills. Therefore, this paper takes the teaching model as one of the core concepts of the study. Teaching mode refers to the following: “it systematically explores the interaction between educational objectives, teaching strategies, curriculum design, and teaching materials, as well as social and psychological theories to model teachers’ behavior in various alternative types of teaching.” The teaching mode is divided into traditional teaching and heuristic teaching in this study. The traditional teaching mode is a teacher-centered, book-centered, and classroom-centered teaching mode. Heuristic teaching is conducted according to the purpose of teaching, content, students’ knowledge level, and the law of knowledge. Using a variety of teaching tools, the inspirational and inducement approach is used to impart knowledge and develop abilities so that students can learn actively to promote physical and mental development.

Early reading refers to “providing infants and toddlers with materials related to visual stimulation and allowing them to receive information about the materials. Based on observation, thinking, and imagination, infants and toddlers can make initial comprehension and verbal expression of the material, express their own opinions, and listen to adult narratives. From 1997 to 2000, the American Council on Early Reading pointed out in its research report, “Preventing Dyslexia in Children at an Early Age,” that there were about 2.56 million school-age children with dyslexia in the United States, accounting for 4.43% of the total school-age population from 6 to 21 [2, 3]. About 80 percent of children with learning difficulties in the United States are dyslexic, and forty percent of U.S. citizens cannot read effectively or even cannot read, which has affected their daily life and the functioning of the social work system. The research report once again put forward the importance of early reading from the perspective of preventing reading difficulties, which has aroused strong response in major educational developed countries in the world. In China, research on early reading education mainly includes the exploration of activities, such as reading materials, reading instruction, and activity evaluation, involved in early reading education in kindergartens and families, and there is still a vague understanding of the basic issues of early reading. Early reading helps children to contact written language, develop their behavior of learning written language, and lay a good foundation for formal written language learning in school age. It is an important way to improve children’s language ability. Early reading can enrich children’s vocabulary, broaden their knowledge, expand the scope of life and learning, form good listening and speaking habits, and help children communicate in appropriate languages. In this study, early reading mainly refers to the general vocabulary mastered by students before entering school.

2. Research Design

2.1. Experiment Groups. The grouping of teaching model is as follows: a total of 22 classes from 7 kindergartens received instructor training. Among them, 11 class is set as the experimental group. The teacher has carried on the experimental training for 2 months. The system theory of heuristic teaching mode, teaching the key core, was carried on during the study and practice.

The control group was composed of the remaining 11 classes. The teachers also participated in the training without learning heuristic teaching, however, they carried out communication and discussion according to the conventional teaching plan to ensure that the teachers in the experimental group would not get experimental intervention, except for the heuristic teaching mode, because of the training. In the middle of the experiment, 1 experimental group and 2 control groups quit the training because of insufficient training time and failed to pass the assessment.

2.2. Survey Subjects. The subject data were selected by the convenience sampling method from 22 classes in 7 public kindergartens in X City from March 1, 2019, to June 1, 2019. Inclusion criteria are as follows: (1) children aged 3 to 5 years, (2) normal IQ, and (3) parents and teachers were informed, and they agreed to the study. Exclusion criteria are as follows: (1) class teacher training fails to reach the standard, and (2) students could not understand the content of the scale. According to the previous analysis, the sample size was 10 times the number of scale items, and the 20% loss to follow-up rate (elimination, shedding, etc.) was considered. Therefore, at least 144 questionnaires should be issued in this study. Finally, 240 questionnaires were sent out and 221 valid questionnaires were collected, with a sample recovery rate of 92.08%. There were 115 males (52.0%) and 106 females (48.0%).

2.3. Scale Design. The hybrid research method, including qualitative research with quantitative research, was used in this study. Firstly, the researcher analyzes and summarizes the concept and dimensions of expressive ability in preschool education using a large number of literatures, reading, and interviews with all teachers involved in training, which are subdivided into four dimensions, namely transmission ability, communication ability, migration ability, and generalization ability. The transmission ability is based on the students’ improvisation, while the generalization ability is based on the teacher’s evaluation of the simplification of students’ retelling stories. Communication ability is to investigate young children involved in communication with teachers’ story elements of the
comprehensive migration ability, which refers to the communication of ideas about the story by involving the children [4].

SPSS25.0 software was used for statistical analysis. Qualitative data were described by frequency and percentage (%). The scale validity was evaluated by content validity and structure validity, which is evaluated by Cronbach’s α coefficient and retest reliability. The critical ratio method and correlation analysis method were used to judge the differentiation and suitability of scale items. The content validity of the scale was evaluated by the expert consultation method, and the structure validity was evaluated by the confirmatory factor analysis. Cronbach’s α coefficient and split-and-half reliability were used to evaluate the reliability of the scale.

The quality of the teacher’s scoring from student data collected in the teacher’s class is assessed to test the reliability and validity of the data obtained using Cronbach’s α coefficient for reliability evaluation. Among them, the total K is for measurement subject, $S_i^2$ is the variation situation of the subject, and $S_x^2$ is all questions of total variance, and they are extracted with the use of the average variance combination reliability to evaluate the questionnaire topic structure validity, where $\lambda$ is the factor load, N is the sample size, and $\epsilon$ is the error term. The specific formulas are as follows:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S_x^2}\right),$$

$$AVE = \frac{\sum \lambda_i^2}{N},$$

$$l_{CR} = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \epsilon}.$$ (1)

The transmission ability evaluated by teachers is on the basis of children’s retelling of the “little white rabbit.” Teachers need to assess three small paragraphs that involve repeat keywords and key elements of scores. In this situation, students communicate with the students in the back row and share the story information they just got from the teacher. Communication ability is composed of students’ expressions observed by teachers in the process of communicating with teachers and other students. Teachers focus on whether students have inappropriate details or unclear themes. The four stories are randomly distributed to students. Students’ transfer ability is based on their situational response. Teachers randomly check whether they can retell a story similar to the one they read. Teachers score students according to their expression of grammar and vocabulary. Considering that different evaluation criteria may lead to unclear distribution of scoring data, the scale and scoring were revised for several rounds to ensure that students’ presentation ability could be measured more accurately. RMSEA and CFI are used to describe the matching degree of the model.

After several rounds of testing of the reliability and validity of items in questionnaire and teacher evaluation criteria, scale and dimensions all passed the standard of reliability and validity. RMSEA and CFI, which represent model fitting, also reach the standard. The concrete numerical value is shown in Figure 1.

2.4. Data Collection. In addition to the children’s expressive ability collected through measurement, this study also collected information on gender, age, family background, and early reading amount of students through a questionnaire survey of parents. The family background based on Bourdieu’s theory of capital is subdivided into three types: economy, culture, and society. Economic capital is composed of parents’ annual income standardized scores. Cultural capital is transformed by parents’ years of schooling. Social capital is classified according to the occupation of parents and occupational standards, and scored by the expert method. Bourdieu believes that children’s family cultural background is long-term. Hence, we should pay attention to the possibility of cultural atmosphere and provide more choices for children. For example, family culture will affect children’s visual cognitive ability and computing ability [5, 6].

A further fitting of family background variables was carried out by AHP hierarchy. It is a system engineering method that regards a complex multiobjective decision-making problem as a whole, decomposes the overall objective into multiple subobjectives or criteria, and then further decomposes into several levels of multiple indicators or criteria and constraints to decompose into an orderly hierarchical structure. Through the method of fuzzy quantification of qualitative indicators to calculate the single ranking (weight number) and total ranking of each level are calculated based on fuzzy quantification for qualitative indicators to optimize the decision-making of multi-objective, multi-index and multi scheme.

The weight vector was fitted by the least square method to minimize the sum of the squares of residuals. The expression is as follows:

$$\min Z = \sum_{i=1}^{n} \sum_{j=1}^{m} (ai_jw_j - w_i)^2 \sqrt{b^2 - 4ac}$$

s.t. $\sum_{j=1}^{m} w_j = 1$

$$w_j > 0, \quad i = 1, 2, \ldots, n.$$ (2)

$W_i$ is the regression coefficient of economic capital, cultural capital, and social capital in the calculation formula of family background, and family background is further expressed as follows:

$$ESCS_i = \beta_0 + W_1 Economic_i + W_2 Culture_i + W_3 Social_i + \epsilon_i.$$ (3)

In addition, since parents cannot have an objective and unified evaluation of students’ early reading amount, the estimation of early reading amount needs to be realized based on certain algorithms. The C4.5 algorithm is adopted
in the decision tree model to do the estimation. The algorithm uses the maximum information gain rate as a measure of node splitting. The use of information gain rate to select attributes can handle incomplete data while avoiding the disadvantage of selecting attributes with too many offset values.

A series of questions on students’ reading behavior were set for parents, and the final reading amount was fitted based on these decision attributes. The decision tree adopts the top-down recursion method, which is a data mining method that classified data through a series of rules. It has a variety of algorithms [7], among which the CHAID algorithm is the chi-squared automatic interaction detector [8]. According to the type of variables, the data was segmented optimally using the chi-square test or variance analysis principle, and the multivariate contingency table was automatically divided according to the $P$-value to generate a multifork tree, which supported discrete target variables and continuous target variables, and the main influencing factor can be efficiently mined.

When a variable is a continuous variable, the CHAID algorithm will determine its grouping according to the analysis of variance from the root node. Each step in the selection and target variable has the strongest effect of input variables. If the category of each input variable has no statistically significant differences between the target variable, it will be merged. When the node data and $P$ value (the significance level, common standard is 0.05) cannot meet the set requirements, the algorithm will stop segmentation.

The attribute values are compared in the internal nodes of the decision tree, and the branch down from this node is judged according to different attribute values. A path in the leaf node of the decision tree corresponds to a jointing rule, and the whole decision tree corresponds to a set of disjunctive expression rules. One of the biggest advantages of the learning algorithm based on the decision tree is that it does not require users to know a lot of background knowledge in the learning process (which is also its biggest disadvantage). As long as training examples can be expressed in the way of attribute-conclusion, the algorithm can be used to learn. This algorithm is briefly described as follows:

$$\text{Gain Information} = \frac{a}{-\sum \frac{|S_j|}{|S|} \log \frac{|S_j|}{|S|}}$$

(4)

The formula above is the calculation of the splitting of decision attributes. In the formula, training set $S$ is divided into $m$ subsets by training math $A$. $|S|$ denotes the size of the sample in the $j$ subdata set, and $|S|$ represents the total sample size. Whether the value of attribute $A$ is the best splitting attribute can be judged by calculating the gain information of attribute $A$ after sample splitting [9]. The calculation formula of gain information is shown in the following formula:

$$\text{Gain Information} (S, A) = E(S) - E_A(S).$$

(5)

Finally, on the basis of these equations, information entropy is further calculated. $E(S)$ is the entropy of training data set $S$, and $E_A(S)$ is the information entropy of attribute $A$ after splitting. The calculation formulas of $E(S)$ and $E_A(S)$ are as follows:
2.5. Evaluation Methods. This study mainly tests the correlation between early reading amount and children’s expressive ability, and it will use independent sample T test, multiple linear regression, and LASSO regression model to explore the relationship between the two. Lasso regression is a least square problem with limited regression coefficients, which can better deal with the “over learning” problem in model learning and realize variable selection while estimating model parameters.

In the evaluation, firstly, the independent sample T-test was used to demonstrate the positive effect of the heuristic teaching model. Then, multiple linear regression was used to calculate the regression coefficient of the early reading amount. Finally, a lasso model was established to further test the correlation between the two.

2.6. Research Procedures. To sum up, in the first round of data sampling, this study first conducted a questionnaire survey among parents, interviewed teachers, generated scales, and conducted reliability and validity tests. It then generated students’ early reading amount in the second round of data sampling and finally generated the final ability variables in the third round of data sampling. The flow chart is shown in Figure 2.

3. Research Results

3.1. Preliminary Difference Analysis. The differences between the groups in early reading, transmission ability, communication ability, migration ability, and generalization ability are shown in Figure 3. Preliminarily, the score of the experimental group in expression ability is significantly better than the control group, and the early reading amount is also higher than the control group.

The independent sample T test was further used to analyze the differences between the two groups and prove the positive significance of the heuristic teaching mode from the data level [10]. The score of the experimental group using the heuristic teaching mode is significantly higher than that of the experimental group using the traditional teaching mode. The difference analysis data are shown in Figure 4. According to the data in the figure, there are significant differences in the four abilities and early reading amount between the two groups of data, among which the difference in transmission is the most obvious.

The difference value is converted into effect value, and the calculation formula is as follows:

\[
\begin{align*}
E(S) &= -\sum_{i=1}^{N} p(x_i) \log_2 p(x_i), \\
E_A(S) &= -\sum_{j=1}^{m} \left| \frac{S_j}{|S|} E(S_j) \right|.
\end{align*}
\]

\(n_1\) and \(n_2\) are the sample sizes of the two groups of data, \(s_1\) and \(s_2\) are the standard deviations of the two groups of samples, and the standard deviations of the samples are obtained after dimensionalization. In the case of mean comparison between the two groups, it is most intuitive to use the mean difference between the two groups as the effect size. However, in psychological studies, when the mean difference of original data is used as the effect size, there will be problems of unit inconsistency, and the effect size cannot be compared between studies. Cohen [11] proposed to replace the original mean difference with the standardized value of the mean difference, which is the basis of the differential effect size. Based on the sample number \(n\) of the experimental group and the control group, the combined standard deviation \(s\) was calculated, and the mean score of the experimental group and the control group was calculated at the same time. The effect size value was calculated using the following formula:

\[
d = \frac{x_1 - x_2}{s}.
\]

There are two interpretations of effect size. One is the relative position (percentile grade) of the experimental group’s mean value in the control group, and the other is the degree of nonoverlap between the distributions of the two groups. Cohen proposed that \(d = 0.2\), \(d = 0.5\), and \(d = 0.8\) correspond to small, medium, and large effect sizes, respectively. The “percentage grade of the experimental group mean in the control group” corresponding to these three effect sizes are 58%, 69%, and 79%, respectively, and the “proportion of nonoverlapping distribution between the two groups” are 14.7%, 33.0%, and 47.4%, respectively. However, Cohen cautioned against blindly using this standard. The effect size of heuristic teaching mode on transmission ability, communication ability, transfer ability, and generalization ability were 0.21, 0.25, 0.31, 0.30, and 0.32, respectively, which were of moderate intensity, indicating that heuristic teaching mode had a good intervention effect on improving students’ expressive ability.

3.2. Regression Prediction under Heuristic Teaching Model. In spite of the positive effects of heuristic teaching model, researchers still hope to establish the relationship between early reading amount and students’ expressive ability. Based on this, the regression model is constructed. Firstly, model (11) is constructed, and only demographic variables and family background variables are included. Ability is student \(i\)’s expressive ability score, and \(Family_1\) is the final fitting family background score. Background is the background variable of the student.

\[
Ability_i = \beta_0 + \beta_1 Background + \beta_2 Family_1 + \epsilon_i.
\]
was determined whether early reading volume had a good explanatory power by observing changes in $R^2$.

\[ \text{Ability}_i = \beta_0 + \beta_1 \text{Background}_i + \beta_2 \text{Family}_i + \beta_3 \text{Reading}_i + \epsilon_i. \]

(10)

Based on this, multiple linear regression was constructed to calculate the regression coefficients of demographic variables and family background variables, respectively, with transmission ability, communication ability, migration ability, and generalization ability as dependent variables, and then, they included the early reading amount. The specific regression coefficients are shown in Figure 5. The bars in the figure show the effect of each variable on the four abilities. Gender has a significant impact on students’ ability of expression. Girls’ ability of expression is 0.12–0.24 standard points higher than boys’ ability of expression. Family background also has a significant impact on students’ ability of expression. The regression coefficients ranged from 0.13 to 0.19.

After the variables of early reading amount were added, the goodness of fit of the model was significantly improved, indicating that early reading amount had a good explanatory power to students’ expressive ability. As can be seen in Figure 6, for each unit increase in early reading, students’ transmission ability, communication ability, migration
ability, and generalization ability will increase by 0.519 points, 0.199 points, 0.204 points, and 0.173 points, all at the level of 0.01, indicating that children’s early reading amount does improve their later expression ability.

3.3. Correlation Test Based on Ridge Regression and Lasso Regression. Considering that the economic, cultural, and social capital of the subordinate dimensions of the family background variable may have different relationships with the early reading amount, it should be differentiated and further controlled. To ensure that the inclusion of such multivariable will not bring collinearity interference, the prediction was carried out using the range regression and lasso regression.

Ridge regression is a linear model commonly used in regression analysis. It can effectively prevent the overfitting of the model. In ridge regression, the model will retain all the characteristic variables but will reduce the weight value of the characteristic variables, and the unified influence of the characteristic variables on the prediction results becomes smaller. This method, which avoids overfitting by retaining all eigenvectors and reducing only the coefficient values of
eigenvectors, is called $L2$ regularization. $L2$ regularization formula is very simple, directly adding the sum of the squares of weight parameters, which is represented as $w_j$ in the subsequent formula, on the basis of the original loss function, and $E_{in}$ is the penalty parameter in the model.

$$L = E_{in} + \lambda \sum_j w_j^2.$$  \hspace{1cm} (11)

Lasso regression is a regularization model for linear regression in addition to ridge regression. Like ridge regression, it also limits the coefficient of feature vector to a range very close to 0, however, it limits the coefficient in a different way. It directly adds the absolute value of weight parameters on the basis of the original loss function. Different from the previous formula, this time, the weight is an absolute value, avoiding excessive fitting, and the specific formula is as follows:

$$L = E_{in} + \lambda \sum_j |w_j|.$$  \hspace{1cm} (12)

In this study, the two methods were used for model fitting, and ROC curves were drawn to compare the goodness of fit of the two models, which can be seen in Figure 7. The comparison showed that the goodness of fit of
Lasso regression was higher than that of ridge regression, and lasso regression kept higher explanatory power in early reading. Therefore, this paper further analyzes the relationship between economic capital, cultural capital, social capital, early reading, and the four expressive abilities, and it presents them in the way of heat map.

According to the heat map (as shown in Figure 8), different amount of early reading has different influence on expressive ability. Specifically, grammar reading has the greatest correlation with communication ability, vocabulary reading has the greatest correlation with transfer ability, and text vocabulary has the greatest correlation with transmission and communication ability. At the level of family background, the correlation of cultural capital is greater than that of other capital, which indicates that cultural capital is one of the variables with profound influence on students’ cognitive ability.

3.4. Mechanism Analysis of Early Reading and Expressive Ability. Finally, the researchers analyzed the possible influence mechanism of early reading quantity from two aspects of dominant influence and recessive influence. The mechanism diagram is shown in Figure 9. First of all, the dominant effects mainly refer to reading and expression ability, and both belong to the category of cognitive dimension. The individual’s early reading can also be regarded as a kind of capital accumulation for the future cognitive behavior [12]. To master certain cognitive ability, the individual makes cognitive investment. The more cognitive capital accumulation in the early stage, the faster the individual can master the cognitive ability of the next stage [13, 14]. In terms of recessive influence, the accumulation of individual reading may not directly promote the improvement of their cognitive ability but will affect the generation of their noncognitive ability. For example, students with more early reading experience may have better learning attitude, confidence, and motivation, which will also obviously improve their expression ability [15, 16].

It should be noted that family plays a “black box” role between these two kinds of influence and expressional ability, and the two kinds of influence are not doomed to promote the improvement of expressional ability. Family may play a role in promoting and improving the positive influence of early reading, however, it may also play a role in eliminating and offsetting a part of the positive influence.

4. Research Conclusions

4.1. Heuristic Teaching Mode Has High Application Value. This study found that compared with other teaching modes, students in the heuristic teaching mode performed significantly better than those in the control group in all aspects, which demonstrated the teaching value of this mode. Heuristic teaching is a dynamic, open, and orderly system. It is the main direction of modern education and teaching reform to constantly emphasize students’ subjectivity. The traditional teaching mode is mostly one-way communication with teachers as the main body, which greatly ignores the enthusiasm and initiative of students in learning. This study demonstrates the precious value of heuristic teaching mode based on the experimental data, which is a more robust piece of evidence.

4.2. There Is a High Correlation between Early Reading Amount and Children’s Expressive Ability. The baby is still in the stage of language development [17], and children’s early reading amount was positively correlated with an expressive ability according to the basic data statistics on all the expressive ability [18]. That is to say, as a general cognitive ability, high reading amount in early stage will account for high expression ability. In contrast, if the amount of early reading is insufficient, children will not be able to achieve the accumulation and precipitation of knowledge, and it will affect the control and play of the language environment and language use. This finding suggests that some early book reading is very important. Teachers or parents should focus on developing children’s interest in reading and helping them to develop some basic reading skills. The only way for children to develop good reading habits is to learn how to read properly.

4.3. The Correlation between Early Reading Amount and Children’s Different Expressive Ability Is Different. For student, the sufficient early reading amount means specific prerequisites for further learning. As a prerequisite, sufficient vocabulary provides efficiency for children to establish
association with novel knowledge. However, this study also found that the facilitating value of early reading amount for four dimensions of expressive ability are not the same. According to Piaget’s theory of cognition [19], language and thought are interdependent, and reading and expression may be the same. However, for each smaller dimension, the mechanism can be complex and entirely different. It means that the impact of early reading on students’ overall ability may be a process of dramatic change.

4.4. Family Background Plays an Important Role. In general, there are two ways that family background influences children’s expressive ability: firstly, families influence children’s expressive ability using their social and economic resources to compete and purchase high-quality educational resources. Secondly, parents cultivate children’s learning interest and habits through educational participation and behavioral support, thus affecting children’s ability to express. In both paths, early reading amount is a potential mediator variable [20, 21].

However, it should be noted that young children’s development is influenced by local sociocultural and family backgrounds. Because of the constraints of the study, the sample selected in this paper did not take into account the influence of geographical factors. Therefore, the findings do not represent the developmental patterns and characteristics of young children across the country. Therefore, the data from other regions should be analyzed in the follow-up study to consider the influence of geographical factors.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This study was supported by Research project of basic Education in Jiangxi Province in 2020: Research on reading assistance for rural preschool children under the guidance of “reading with adults” (No. SZUGGYW2020-965).

References

[1] I. E. Boerma, S. E. Mol, and J. Jolles, "The role of home literacy environment, mentalizing, expressive verbal ability, and print exposure in third and fourth graders’ reading comprehension," Scientific Studies of Reading, vol. 21, no. 3, pp. 179–193, 2017.
[2] I. L. V. V. D. Spek, M. C. J. P. Franken, R. M. Swarte, and N. W. Kuperus, "Validity of an early parent-report questionnaire for language disorder in very preterm children from 2 to 10 years of age," European Journal of Paediatric Neurology, vol. 34, pp. 1–6, 2021.
[3] T. V. Phan, D. M. Sima, C. Beelen, J. Vanderauwera, D. Smeets, and M. Vandermosten, "Evaluation of methods for volumetric analysis of pediatric brain data: the child metrix pipeline versus adult-based approaches," Neuroimage Clinical, vol. 19, pp. 734–744, 2018.
[4] R. W. Cheung, C. Hartley, and P. Monaghan, "Receptive and expressive language ability differentially support symbolic understanding over time: p," Journal of Experimental Child Psychology, vol. 214, Article ID 105305, 2022.
[5] N. Notten and B. Be Cker, "Early home literacy and adolescents’ online reading behavior in comparative perspective," International Journal of Comparative Sociology, vol. 58, no. 6, pp. 475–493, 2017.
[6] A. Breinholt and M. M. Jger, "How does cultural capital affect educational performance: signals or skills?" British Journal of Sociology, vol. 71, 2020.
[7] B. Talekar, "A detailed review on decision tree and random forest," Bioscience Biotechnology Research Communications, vol. 13, no. 14, pp. 245–248, 2020.
[8] Y. Xia, C. Liu, Y. Y. Li, and N. Liu, "A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring," Expert Systems with Applications, vol. 78, pp. 225–241, 2017.
[9] G. Ke, M. Qi, F. Thomas et al., "LightGBM: a highly efficient gradient boosting decision tree," in Proceedings of the Neural Information Processing Systems, Curran Associates Inc, Long Beach, CA, U.S.A, December 2017.
[10] R. A. Cooper, "Making decisions with data: comparing sun leaves to shade leaves with a t-test," The American Biology Teacher, vol. 83, no. 4, pp. 229–234, 2021.
[11] J. Cohen, The Effect Size Index: D. Statistical Power Analysis for the Behavioral Sciences, Routledge, New York, NY, U.S.A, 1988.
[12] A. Strobel, G. Wieder, P. C. Paulus et al., "Dispositional cognitive effort investment and behavioral demand avoidance: are they related?" PLoS One, vol. 15, no. 10, 2020.
[13] B. S. Nayak and P. K. Sahu, "Socio-demographic and educational factors associated with depression, anxiety and stress among health professions students," Psychology Health & Medicine, vol. 27, no. 4, pp. 848–853, 2021.
[14] M. P. Ferra, R. L. Quijano, and I. G. Martínez, "Impact of educational habits on the learning of 3–6 Year old children from the perspective of early childhood education teachers," Sustainability, vol. 12, no. 11, p. 4388, 2020.
[15] Y. Wang, Z. Yang, Y. Zhang, F. Wang, T. Liu, and T. Xin, "The effect of social-emotional competency on child development in western China," Frontiers in Psychology, vol. 10, p. 1282, 2019.
[16] G. M. Pion and M. W. Lipsey, "Impact of the Tennessee voluntary prekindergarten program on children’s literacy, language, and mathematics skills: results from a regression-discontinuity design," AERA Open, vol. 7, no. 2, pp. 233285842110413–233285842110497, 2021.
[17] I. Saitadze, "Mediating effects of early childhood programs and high quality home environments on the cognitive development of poor children involved in the child welfare system," Children and Youth Services Review, vol. 120, Article ID 105736, 2021.
[18] Z. Xu and D. Liu, "A comparison of embodied methods to improve Chinese children’s reading comprehension: observed and participant performed manipulations," Journal of Research in Reading, vol. 43, no. 4, pp. 556–576, 2020.
[19] L. Boonk, H. J. Gijselaers, H. Ritzen, and S. B. Gruwel, "A review of the relationship between parental involvement
indicators and academic achievement,” *Educational Research Review*, vol. 24, pp. 10–30, 2018.

[20] L. Wmann, “Educational production in east asia: the impact of family background and schooling policies on student performance,” *German Economic Review*, vol. 6, no. 3, pp. 331–353, 2010.

[21] W. Tomaszewski, N. Xiang, and M. Western, “Student engagement as a mediator of the effects of socio-economic status on academic performance among secondary school students in Australia,” *British Educational Research Journal*, vol. 46, no. 3, pp. 610–630, 2020.