Using Learning Analytics to Monitor and Enhance Early Reading Fluency: A Study in Brazilian Public Schools

Thomaz Edson Veloso da Silva[1], Guilherme Corredato Guerino[2], Luiz Rodrigues[3]

[1] Educometrika LTDA
Fortaleza – CE – Brazil

[2]State University of Paraná (UNESPAR)
Apucarana – PR – Brazil

[3]Federal Technological University of Paraná (UTFPR)
Apucarana – PR – Brazil

thomaz@educometrika.com.br, guilherme.guerino@ies.unespar.edu.br

luizrodrigues@utfpr.edu.br

Abstract. *This paper examines the use of Learning Analytics (LA) to monitor and improve early reading fluency in Brazilian public schools. In July 2024, the Chapecó city launched a policy for assessing 2,072 second-grade students over six monthly cycles using three distinct instruments. Students were classified into six fluency categories, enabling targeted interventions. Results showed a substantial reduction in the number of Non-readers and a significant increase in Fluent Readers across six applications. These findings suggest that systematic application of LA tools, combined with regular teacher feedback and data-informed instructional adjustments, can produce measurable improvements in early literacy outcomes, offering insights for policymakers and educators seeking to implement data-informed strategies in early reading programs.*

1. Introduction

Developing reading fluency in the early grades is a critical factor influencing future academic success and social equity [Panel 2000]. In Brazil, despite various policy efforts, a significant portion of second-grade students complete this stage without acquiring adequate reading fluency, hindering their ability to access more advanced learning opportunities [Camillo 2024].

Learning Analytics (LA) has emerged as a promising approach to support educational improvement by collecting, analyzing, and communicating data about learners and their learning processes, thereby informing teaching practices and enhancing student outcomes [Siemens 2013]. In July 2024, the municipal education department in Chapecó launched a policy initiative leveraging LA to systematically assess and improve early reading fluency, inspired by the Early Grade Reading Assessment (EGRA) framework [Gove and Wetterberg 2011].

The initiative monitored 2,072 second-grade students through six monthly applications of three distinct instruments measuring different aspects of reading development. By applying LA methods to this data, the program aimed to provide actionable insights for teachers and school administrators, fostering targeted pedagogical interventions and

DOI: 10.5753/wapla.2025.15887

continuous improvement cycles. This paper presents the context, methodology, results, and implications of this initiative, offering evidence on the practical application of LA to address foundational literacy challenges in Brazilian public schools. By presenting these insights, this paper contributes to the LA community with evidence on how systematic, well-planned assessments can inform educational interventions, culminating in valuable learning gains and, ultimately, improving learners' development.

2. Method

This paper reports a case study concerning the collection and usage of LA in a city-wide educational systems. The focus is on young learners' development of reading fluency, presenting data collected over a full semester. To describe this study, the remainder of this section introduces its context, participants, assessment instruments, reading fluency classification, and data analysis.

2.1. Context

In July 2024, the municipal education department of Chapecó initiated a policy focused on systematically assessing early reading fluency, inspired by the EGRA framework [Gove and Wetterberg 2011]. The initiative aimed to provide real-time, actionable data using LA tools, supporting targeted pedagogical interventions to improve literacy outcomes among second-grade students in the public education system.

2.2. Participants

A total of 2,072 second-grade students from public schools in Chapecó participated in the initiative. Data was collected monthly over six consecutive applications, from July to December 2024. Importantly, this paper reports aggregated data, gently provided to us by the municipal education department of Chapecó. On the one hand, given that this kind of data does not allow the identification of particular participants, it does not require this study to be reviewed by an Institutional Review Board. On the other hand, it limits our ability to describe our sample in more detail. Nevertheless, to cope with this limitation, we rely Brazilian goverment's data. The age-grade distortion in Chapecó was 5.4% in 2024 for early grades[1]. Accordingly, it is likely that most students from our sample were six years old at the time of data collection.

2.3. Assessment Instruments

Following the EGRA framework [Gove and Wetterberg 2011], three distinct instruments were used in each assessment cycle:

1. **Oral Comprehension:** the evaluator read aloud a simple text and posed three comprehension questions to the student: two questions required detecting explicit information, and one required inferring implicit information.
2. **Word List:** each student was shown a chart with 60 words of increasing complexity, progressing from canonical to non-canonical forms. The evaluator recorded the number of words the student read correctly within a one-minute period. The fluency benchmark was set at 45 words per minute.
3. **Text Reading and Comprehension:** students read a simple text aloud and then answered three comprehension questions, following the same structure as the oral comprehension task.

[1]https://qedu.org.br/municipio/4204202-chapeco

2.4. Classification

According to the EGRA framework and toolkit [International 2016], classifying students into multiple reading development categories helps identify the continuum of skills that learners develop on the path toward fluent reading. This classification enables educators and policymakers to tailor interventions according to distinct stages of reading acquisition, from those unable to decode text to those who read fluently with comprehension. Such detailed categorization provides actionable diagnostic data to guide instructional planning and monitor progress effectively [International 2016]. Therefore, based on the student's performance measured by the instruments presented before, they would be classified in one of the following stages:

- **Non-reader:** Unable to identify letters, syllables, or words; shows no or minimal decoding ability.
- **Syllable reader:** Can decode and read syllables but struggles with whole words.
- **Word reader:** Can read familiar words accurately but with limited fluency or comprehension.
- **Phrase reader:** Reads short phrases and simple sentences with some fluency and basic understanding.
- **Fluent reader:** Reads connected text smoothly and accurately with good comprehension.

2.5. Data Analysis

Results from each application were compiled into a digital database. LA techniques were applied to track the progression of each student and generate aggregate reports on the distribution of students across fluency categories. These insights were shared with educators through dashboards and reports, supporting pedagogical planning and instructional adjustments throughout the six months.

Moreover, while reiterate that this paper's results are based on aggregated data. As a result, we cannot perform comprehensive analysis like trend-based regressions and derive robust inferences from our data. However, we can perform inferential tests aligned with our data structure (i.e., count-based data). Therefore, we perform two-proportion z-tests, which is commonly used to compare proportions between two groups, to understand changes in the number of students classified within a given reading fluency over time.

3. Results

Over the six-month assessment period, significant shifts were observed in the distribution of students across the six reading fluency categories. Figure 1 highlights a positive trend, in terms of reading classification by month, demonstrating that several students improved their reading fluency after each assessment period. Overall, these insights suggest that the LA-based approach contributed to enhancing students reading fluency.

Furthermore, Figure 2 presents a bar plot concerning reading fluency categories and assessment periods. Through a complimentary perspective, this figure shows the changes of student counts in each category over the six assessment applications. Based on that, the remainder of this section highlights some notable insights into how students reading fluency levels changed over time.

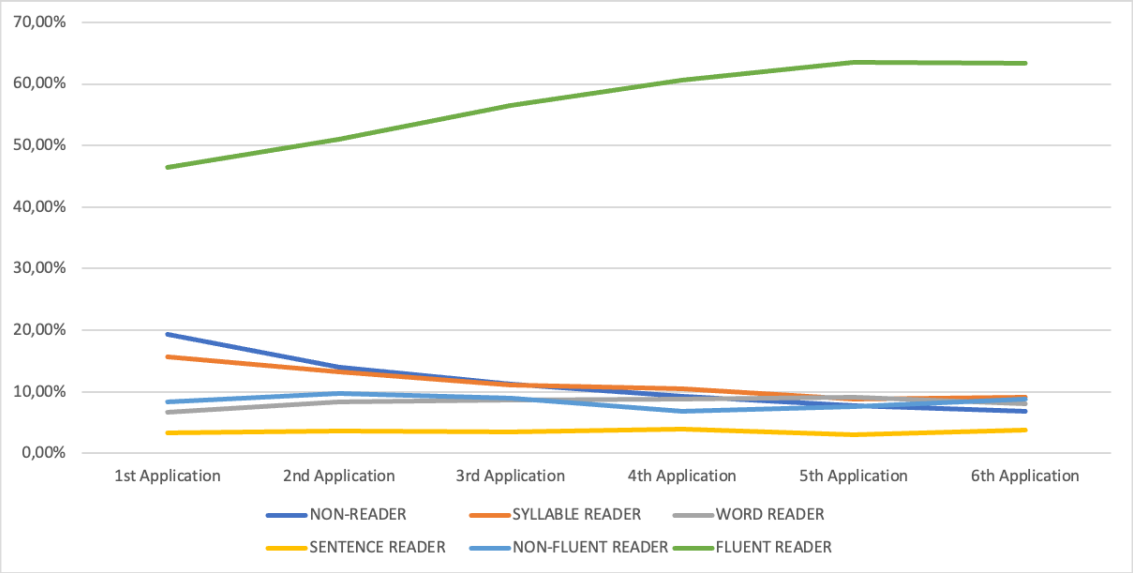


Figure 1. Line chart presenting the progression of student fluency categories over each period of assessment.

3.1. Reduction in Non-Readers

At the beginning of the assessment cycle, 395 students (approximately 19% of the total) were classified as *Non-readers*. By the sixth application, this number had dropped to 143 students (approximately 7%), representing a reduction of 252 students. According to the z-test of two proportions, this difference is statistically significant at the 95% confidence level (χ = 129.7; DF = 1; $p < 0.001$; CI = [0.099 – 0.141]).

3.2. Increase in Fluent Readers

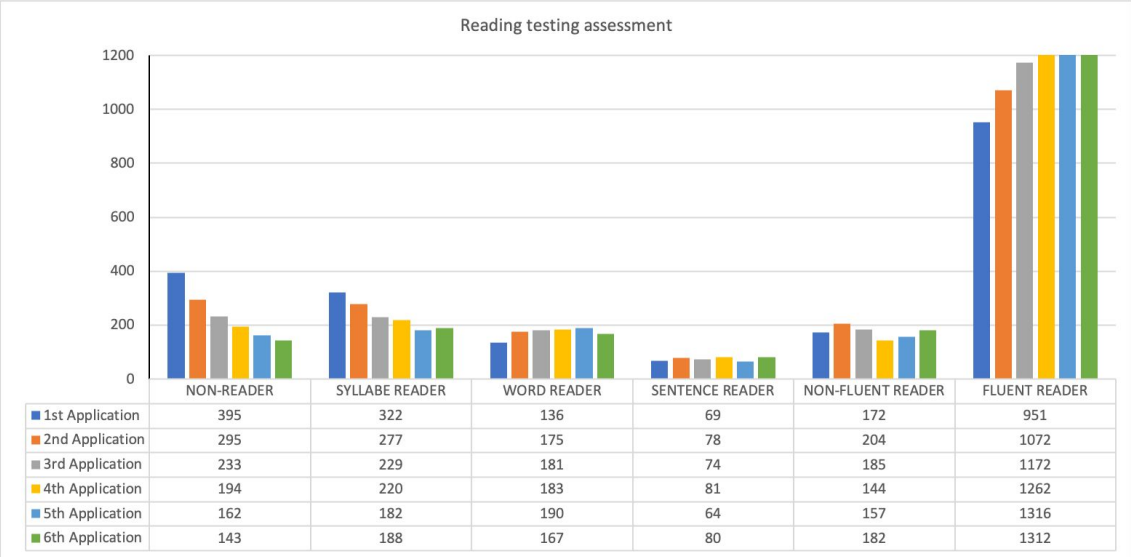
Conversely, the number of students classified as *Fluent Readers* increased from 951 in the first application (46%) to 1,312 in the sixth application (63%), reflecting an improvement of 361 students. According to the z-test of two proportions, this difference is statistically significant at the 95% confidence level (χ = 119.92; degrees of freedom = 1; $p < 0.001$; CI = [-0.200 – -0.139]).

3.3. Overall Trends

The data indicate steady reductions in the *Non-reader* and *Syllable Reader* categories, as well as continuous growth in the *Fluent Reader* category, across all six applications. These trends suggest that the systematic use of LA, combined with teacher feedback and targeted intervention, contributed to notable gains in early reading fluency among participating students.

4. Discussion

The results demonstrate the potential of Learning Analytics (LA) to influence early literacy outcomes in large-scale public education settings positively. Over six monthly assessment cycles, the number of *Non-readers* decreased by 252 students, while the number of *Fluent Readers* increased by 361 students. These improvements suggest that systematic monitoring and feedback loops, enabled by LA, provided valuable information to educators, allowing for targeted instructional adjustments.



	NON-READER	SYLLABE READER	WORD READER	SENTENCE READER	NON-FLUENT READER	FLUENT READER
1st Application	395	322	136	69	172	951
2nd Application	295	277	175	78	204	1072
3rd Application	233	229	181	74	185	1172
4th Application	194	220	183	81	144	1262
5th Application	162	182	190	64	157	1316
6th Application	143	188	167	80	182	1312

Figure 2. Progression of student fluency categories across six applications.

The continuous reduction in lower categories (*Non-readers* and *Syllable Readers*) and the corresponding growth in *Fluent Readers* reflect how providing teachers with data-driven insights can support differentiated instruction and more effective intervention strategies. Prior studies have highlighted that teacher access to real-time data contributes to more responsive pedagogical practices and improved student outcomes [Siemens 2013, Cunha and Capellini 2017].

Despite the promising results, challenges emerged, including variability in evaluator training, occasional delays in data entry, and resistance from some teachers unfamiliar with data-driven approaches. Addressing these challenges will be crucial for sustaining gains and scaling similar initiatives in other municipalities.

Furthermore, while this study tracked fluency progression quantitatively, future work should also investigate qualitative dimensions of student engagement and the fidelity of instructional interventions informed by LA.

5. Conclusion

This study examined the application of Learning Analytics (LA) in supporting early reading fluency development in Chapecó, Brazil. Over the course of six months and six assessment cycles, 2,072 second-grade students were monitored using instruments inspired by the Early Grade Reading Assessment (EGRA). Results showed a substantial reduction in the number of *Non-readers* and an increase in *Fluent Readers*, highlighting the value of using LA to inform and improve pedagogical practices.

The initiative demonstrated that systematic data collection and analysis can provide educators with actionable insights, enabling them to implement targeted interventions and deliver differentiated instruction. Challenges such as variability in evaluator practices and teacher familiarity with data tools underscore the need for continued professional development and support.

Future research should explore the long-term retention of fluency gains, the quali-

tative aspects of instructional practices informed by LA, and how similar frameworks can be adapted to different regional and linguistic contexts. Overall, this study contributes evidence supporting the role of LA in enhancing foundational literacy and guiding educational policy.

Acknowledgments

We acknowledge the use of Generative AI to help in drafting and revising this paper, which was revised accordingly, and take full responsibility for its content.

Ethical Concerns

This study relies on aggregated data, previous collected and gently provided to us by the municipal education department of Chapecó. Accordingly, it is not eligible for assessment by an Institutional Review Board.

Artifacts Availability

This study's artifacts will be made available from the corresponding authors.

References

Camillo, E. d. S. (2024). Social sustainability and reading promotion: linkages between public policies and quality education in brazil. *Em Questão*, 30:e–135828.

Cunha, V. L. O. and Capellini, S. A. (2017). Informative intervention programs to reading comprehension: Development and implementation. *Estudos de Psicologia (Campinas)*, 34(3):411–422.

Gove, A. and Wetterberg, A. (2011). *Early Reading: Igniting Education for All. A report by the Early Grade Reading Assessment (EGRA) program*. RTI Press.

International, R. (2016). *Early Grade Reading Assessment (EGRA) Toolkit, Second Edition*. RTI Press, Washington, DC.

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

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10):1380–1400.