LXD RESEARCH
ESSA
Level 2
Evidence

LEARNING EXPERIENCE DESIGN | LXD Research | engage2learn

Exploration of ELL Student WIDA Access Growth

GRADES 1-5, 20-21, 21-22, & 22-23 SCHOOL YEARS

Authors

Rachel Schechter, Ph.D.

Paul Chase, Ph.D.

Anna Robinson, M.A.

December 14, 2023

engage2learn

Learning Experience Design (LXD)
Research & Consulting
A division of Charles River Media, Inc.

Abstract

This study explores the implementation of the engage2Learn program among English as a Second Language (ESL) educators in the Providence Public School District and its impact on student outcomes over three years (2021-2023). Utilizing Grade 1-5 WIDA ACCESS scores, this study examines cohort differences in change in student proficiency levels before and after e2L coaching. The results indicated positive associations between e2L coaching and ELL student ELA proficiency as indexed by year-over-year growth in WIDA Access scores. In addition to the overall effect of e2L coaching, ELL student growth was greatest among students of teachers who earned two or more e2L best practice badges. Qualitative data in the form of teacher surveys supported the quantitative findings with positive feedback from participating educators. These educators also noted that they had actively integrated e2L coaching strategies into their daily teaching practices. The study suggests a significant benefit when ELL teachers build capacity in educational best practices aligning with high-quality ESL program characteristics. Despite some limitations including non-random assignment of teachers, the study provides insights for educator professional development, ESL policymakers, and educators.

Table of Contents

Introduction

The mismatch between student needs and qualified English Language Learning support educators is substantial. According to the U.S. Department of Education, 9% of all students, nearly four million public school students are English as a Second Language (ESL[1]) learners (NCES, 2021). Unfortunately, only nine percent of fourth-grade ESL students were at or above proficient in reading on the 2017 National Assessment of Educational Progress, and in eighth grade, only five percent were proficient (NCES, 2018). Nationally, after 2020, ESL growth declined by six scale score points in the 1–2 grade-level cluster and by nine scale score points in the 3–5 grade-level cluster (WIDA Research Report, 2021). Specifically, growth in Speaking was substantially lower among all grades.

In 2020-2021, student performance may have been impacted due to disruptions related to the COVID-19 pandemic. Several WIDA Consortium member states reported irregularities in testing windows, with some students testing earlier in the school year, which would shorten opportunities for instruction. In contrast, other students were tested at the end of summer. Indeed, higher proficiency and growth were observed among students who tested later in the school year (Sahakyan & Cook, 2021). According to results from 35 states that administer the WIDA ACCESS assessment, overall, average composite English-language-proficiency scores trended lower in the 2021-22 school year than in pre-pandemic 2018-19 and 2019-20 years, with particular declines in the younger elementary grades (Najarro, 2023).

English learners come from diverse language backgrounds. Differences among this population can range widely based on previous schooling, socioeconomic status of their families, age, first language, and cultural origin. As such, their English proficiency levels can vary significantly across multiple domains. Some students may have a strong command of English by 5th grade through exposure to traditional English Language Arts and Reading core programs, particularly in culturally inclusive classrooms, while others may still be in the early stages of language acquisition at the end of elementary school (Abedi, 2007; Talandis & Stout, 2015). This variation can impact their math and reading assessment performance and long-term academic success. Federal law requires that all states receiving federal funds implement school-based programming based on English Language Proficiency (ELP) standards designed to raise proficiency in academic content and academic English. Schools must also actively monitor students' English proficiency progress and report annually using assessments such as WIDA ACCESS for English learners, the prominent tool used in U.S. schools (Lee, 2012).

While the length of exposure to English instruction impacts students' language acquisition rate, the effectiveness of instruction plays a crucial role. Schools with strong ESL (English as a Second Language) programs rooted in research-based instructional practices and dedicated language support for ELs tend to produce better outcomes (Stepanek et al., 2010). ESL education often falls short due

[1] MLL, ELL and ESL are used interchangeably to refer to students who are learning English and speak a different first language. When speaking about research conducted in this space, the language of the authors or MLL will be used.

to the absence of standardized programs, clearly defined student expectations, and a shortage of adequately trained teachers (Zen, 2001). There are several strategies that districts and schools can implement to ensure high-quality ESL instruction: 1) Make success for ESL students a central issue; 2) Build bridges throughout the community; 3) Use consistent and reliable practices and procedures to identify, assess, and place ESL students; 4) Implement research-based strategies; and 5) Choose appropriate ESL program models that are understood by all staff members (Stepanek et al., 2010).

By building the skills and capacity of educators who work with ESL students, schools will undoubtedly increase English learners' overall proficiency. Only two percent of public school teachers are ESL instructors, and fewer than half of all teachers have taken at least one course on how to teach ESL students (NCES, 2019). With only one certified instructor for every 150 ESL students and more than half of all teachers without any training, there is an urgent need for effective, engaging tools that are helpful to all students (including ESL students) and that are easily used by all teachers. Furthermore, high-quality ESL education should prioritize the holistic development of students, not only teaching them the language but also acknowledging and supporting their experiences, emotions, and culture (Van Hgo, 2007).

Unlike traditional professional development programs, engage2learn seeks to empower individuals through coaching that connects with each learner's cultural and lived experiences. Since 2011, engage2learn has partnered with public schools in over 300 districts, serving 79,000 educators and 4.5 million students across the United States. Professional development has traditionally been a compliance-based, one-size-fits-all model that too often leaves public school systems without the tools, insights, or platform to address the challenges facing educators today. Engage2learn is committed to shifting and scaling how public schools support educators using the Talent Development Pillars to transform PD with professional learning solutions and their all-in-one coaching platform to bring out the best in every teacher and their students. English learners whose teachers were not only coached using engage2learn but achieved multiple badges saw year-over-year gains in Overall Scale Score, Literacy, and Writing.

There is a critical opportunity to use technology better to support our educators of ESL students of all ages, to support job retention and advancement, and to enhance educator best practices that address variable learner needs. Within this context and informed by this need, engage2learn (e2L) partnered with LXD Research to conduct a third-party evaluation of the effects of e2L implementation during the Spring and fall of 2022 in the Providence Public School District (PPSD) in Rhode Island.

Method

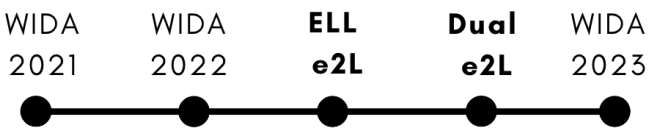
Research Description

The primary goal of this research study is to examine the impact of e2L coaching of MLL teachers on English Language learning. Additional goals of the study are to examine specific effects of levels of engagement with e2L's all-in-one instructional coaching platform, GroweLab, as indexed through the earning of best practice badges. To meet these goals, LXD Research designed a longitudinal study of e2L to start understanding this program's impact on student language achievement in PPSD.

The Providence Public School District is located in Rhode Island, with approximately 22,000 students in grades K-12 across 37 schools in the district. The district serves a population of 16% of students receiving special education services, and nearly 33% are Multi-Language Learners (MLL). Approximately 65% of the students served are Hispanic, 16% Black, 9% White, 5% Asian, 4% Multi-racial, and 1% Native American. Among the student population, 55 native languages are spoken, and 91 countries are represented.

LXD Research conducted a quasi-experimental study that followed students over multiple years using annual WIDA ACCESS standardized tests. While Providence MLL educators received coaching during the 2021-2022 school year, the WIDA ACCESS English proficiency assessment was given yearly in the winter months. Given the training and assessment timing, this study focused on student scores from Winter 2021 through Winter 2023, which would speak to the impact of coaching during Spring 2022. Figure 1 presents the timeline of the assessments and interventions. MLL teachers were coached in the first cohort. While some Dual Language (Dual) teachers were coached alongside the MLL teachers, additional Dual Language teachers were coached in a second cohort.

Figure 1. Assessment and Intervention Timeline



Research Questions

The following research questions were used to guide our analyses. They are split into:

1. Educator Usage and Impact question, and
2. Student Impact questions

RQ1: Educator Usage and Impact Questions

- How many educators have received coaching, and when?

- How many badges have educators earned through engage2learn, and in which instructional Best Practices (BPs)?
- To what extent does coaching impact educator retention?

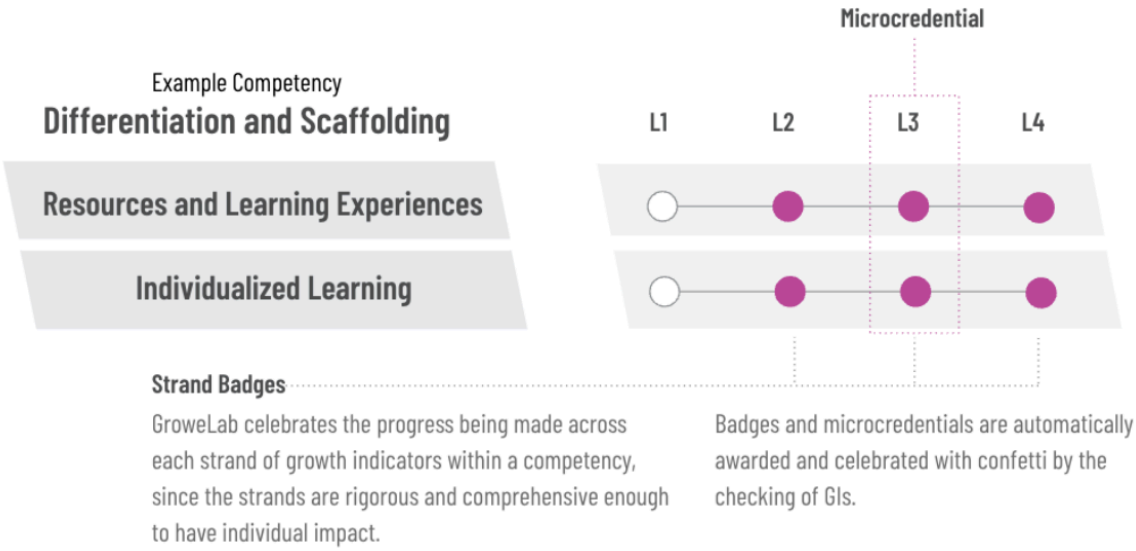
RQ2: Student Impact Questions

- How does an educator's participation in coaching through engage2learn impact the percentage of their MLL students who improve in relative proficiency overall and on each of WIDA's subtests?
- How does an educator's participation in coaching through engage2learn impact their MLL students' improvement on Overall Scale Scores and each of WIDA's subtests?
- How does an educator's attainment of badges (i.e., which types of e2L Best Practice badges, and how many) impact their MLL students' level of improvement on Overall Scale Scores and each of WIDA's subtests?
- How does student growth in the e2L-coached teacher sample compare to Access growth in the general population of Providence students with Access data?
- Specifically, how are the students of the dual-language teachers performing with the extra semester of training?

Engage2learn GroweLab Program Description

Participating teachers were coached and received online resources for a series of educational Best Practices. Highly engaged teachers can earn e2L best practice microcredentials, indicating at least three Growth Indicators (GIs) across a set of related Best Practice competencies. Progress for each teacher and all related materials and resources are maintained in GroweLab, e2L's all-in-one teacher support and instructional coaching platform.

Figure 1. Example Best Practices



Note: The above image represents the system for Microcredentials as of Summer 2023, immediately following the conclusion of data collection. This study includes L1 GI attainment as badge attainment.

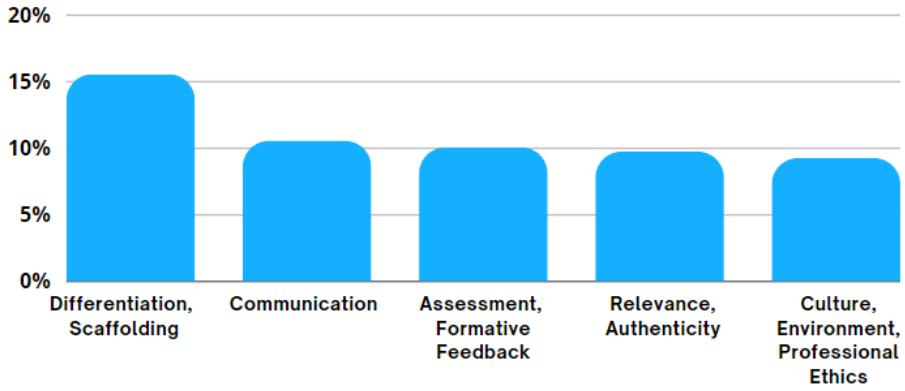
Engage2learn GroweLab Coaching data

Engage2learn coaches and coachees follow a customizable coaching timeline with clearly defined growth benchmarks. This coaching timeline sequences growth on professional learning standards founded on local and state requirements and research-based best practices. The coaching timeline standardizes the coaching experience across the organization while remaining customizable to each coachee's needs. Coachees can choose goals most relevant to their professional learning objectives. The coaches use the GroweLab online coaching facilitation and reporting system to document their coaching efforts and the teachers' progress in earning best practice badges. Therefore, engage2learn provided LXD Research with user implementation data via GroweLab to determine which educators in a school were coached and any differences in their level of improvement in best practices over time.

Teacher Usage and Outcomes

Across the district and multiple educator roles, 408 Providence teachers earned 1,758 best practice badges. All 12 e2L Best Practices were represented and were relatively evenly distributed. The most commonly earned badges were Differentiation & Scaffolding, Communication, and Assessment & Formative feedback (Figure 2).

Figure 2. Top Five e2L Best Practice Badges Earned by Providence Teachers



Teacher Survey

A Teacher Survey was administered to get feedback for engage2learn and hear teachers' voices to complement the student data story. The district administered it online and sent it to the MLL and DLL instructors. There were 12 respondents who completed the survey all the way through and one respondent who completed half the survey.

Student Language Proficiency Assessment

WIDA ACCESS provides Scale Scores for seven different areas of English language proficiency to support monitoring the growth of student skills year over year within each domain. Scale Scores consider item difficulty, so educators can use them to examine groups of students or student performances over time. Scale Scores are specifically designed to track students' growth over multiple years. Proficiency levels are always calculated from the Scale Scores and are grade-level specific. For these reasons, this analysis focuses on year-over-year growth using scale scores across grades 1-5. However, we also use Proficiency scores to compare the likelihood of improvement for students pre- and post-e2L instructor coaching.

Table 1. Example Rubrics of WIDA Domains

Students who score high on **LISTENING** can understand oral language in English and participate in all academic classes and:

- Synthesize information from multiple speakers
- Recognize language that conveys information with precision and accuracy
- Create models or visuals to represent detailed information presented orally, and
- Identify strengths and limitations of different points of view.

Students who score high on **SPEAKING** can use English to communicate orally and participate in all academic classes and:

- React and respond to multiple points of view,
- Organize and present research-based information,
- Clarify how or why something happens, and
- Persuade others based on opinions, examples, and reasons.

Students who score high on **READING** can understand written language in English from all academic classes and:

- Evaluate written information from various sources of information,
- Conduct research and synthesize information from multiple sources,
- Distinguish various processes based on details in written texts, and
- Recognize different ideas, claims, and evidence about a variety of issues.

Students who score high on **WRITING** can communicate in writing in English using language from all academic classes and:

- Produce organized commentaries and editorials on various issues,
- Elaborate narratives with rich, descriptive language and complex organization,
- Create formal written reports on a variety of issues, ideas, and information, and
- Produce well-organized persuasive essays using complex and technical language.

In addition to proficiency level and Scale Scores for each language domain, students receive a proficiency level score and a Scale Score for different combinations of the language domains. These composite scores are Oral Language (50% Listening, 50% Writing), Literacy (50% Reading, 50% Writing), Comprehension (70% Reading, 30% Listening), and Overall (35% Reading, 35% Writing, 15% Listening, 15% Speaking).

Student Sample Selection

As part of the study, PPSD provided LXD Research with student demographic information and achievement information related to English Language learning (i.e., WIDA ACCESS data) administered during the Winter of the 2021, 2022, and 2023 school years. Students are given WIDA ACCESS every winter and end-of-year tests for its main subjects. Out of 10,332 students in PPSD who were given the WIDA ACCESS from 2021-2023, 6,525 students were identified as having taken the test at least two years in a row.

To assess the effectiveness of e2L coaching on student WIDA ACCESS performance, LXD Research identified instructors who taught during both 2021-22 and 2022-23 school years using student class roster data provided by PPSD school administrators and matched those instructors with students who had WIDA ACCESS data for two consecutive years. Teachers who started earning badges before Spring 2022 and those who earned badges after the 2023 WIDA ACCESS test were excluded from the analysis. The resulting sample was 38 elementary school MLL teachers who taught during both years and started earning badges in the Spring of 2022. Because students took the WIDA

ACCESS test in the winter term, we can say that these teachers were effectively uncoached before their students took the Winter 2022 test.

The resulting sample of students was split into two groups (Table 2). The first group of students includes those who had WIDA ACCESS scores for 2021 and 2022 and were taught by teachers pre-e2L coaching during the 2021-2022 school year. The second group includes students who had WIDA ACCESS scores for both Winter 2022 and Winter 2023 and were taught by the same teachers post-e2L coaching. Some students (n = 121) were included in both groups as they had scores for all three years and were taught by the same teachers.

Students with scores assessed between 2021-2022 will be considered the Pre-e2L Coaching group and the 2022-2023 students will be the Post-e2L Coaching group. Groups did not significantly differ by age, gender, race, ethnicity, time spent in MLL programs, or baseline overall WIDA ACCESS Scale Score. Differences between these two groups on WIDA ACCESS will point to the impact of e2L coaching.

Table 2. Elementary Education Student Sample

Cohort	School Year	Grade					Total Students	Demographics
		1st	2nd	3rd	4th	5th		
Pre-e2L coaching	2021-2022	114	86	167	64	96	527	92% Hispanic/Latino 56% Female
Post-e2L coaching	2022-2023	119	66	133	79	72	469	

Exclusion Criteria

There were students who didn't fit the above criteria but were still matched with e2L-coached teachers. They fell into the following categories and were excluded from the main analysis:

- Students with two test scores (2022 and 2023), but were taught by teachers pre-e2L coaching the previous year (2021-2022) and by teachers that were not coached through e2L in the 2022-2023 school year. While they took the WIDA ACCESS in Winter 2022 while being taught by a teacher pre-e2L coaching, any gains in Scale Scores would be associated with teachers not coached through e2L.
- Students with 2 test scores (2021 and 2022) but were taught by teachers who were not coached through e2L during that time and weren't matched with teachers coached through e2L until 2022-2023. Since they did not take the WIDA ACCESS in 2023, there is no score to assess the impact of e2L coaching.

- Students with only one WIDA ACCESS score that can't be included in a change-over-time analysis.
- Students who have been matched with teachers but have no ACCESS data.

Student Language Proficiency Results

Longitudinal Analysis of Pre- and Post-e2L Coached MLL Teachers

Higher Likelihood of Growth in Proficiency

Engage2learn coaching had a positive effect on the likelihood of students improving in proficiency level (Tables 3-4). Students taught by teachers post-e2L coaching showed a significantly higher likelihood of experiencing an overall increase in proficiency than the students of these teachers before participating in the program. This overall difference was also individually significant when analyzing second[2] and third-grade[3] students (Figures 3-4). Furthermore, when considering composite proficiency scores, students were more likely to demonstrate improvement in **Literacy** and **Comprehension** proficiency. When assessing domain-specific proficiency, there was a significant impact on **Writing**, particularly evident in Grades 3 and 4 (Figures 4-5). Additional subgroup analyses for this measure are located in the appendix.

Table 3. Student Outcomes by Percent Likelihood of Improvement in Proficiency by Domain

Instructor Coaching	Overall Score	Oral Language	Literacy	Comprehension
Pre-e2L	**78%**	68%	**74%**	**66%**
Post-e2L	**83%***	72%	**80%***	**74%***
T-Statistic	**2.28**	1.41	**2.42**	**2.69**
Effect Size	**.14**	n/a	**.15**	**.17**

Note: *p<.05, **p<.01, ***p<.001.

[2] t(144) = 2.60, p<.05, Cohen's d (Effect Size) = .42
[3] t(289) = 2.92, p<.01, Cohen's d (Effect Size) = .34

Table 4. Student Outcomes by Percent Likelihood of Improvement in Proficiency by Subscale

Instructor Coaching	Listening	Speaking	Reading	Writing
Pre-e2L	70	66	61	**71**
Post-e2L	71	70	68	**82*** **
T-Statistic	.14	1.06	2.24	**4.35**
Effect Size	n/a	n/a	n/a	**.28**

Note: *p<.05, **p<.01, ***p<.001.

Figure 3. Significantly Higher Likelihood of Proficiency Improvement on WIDA ACCESS

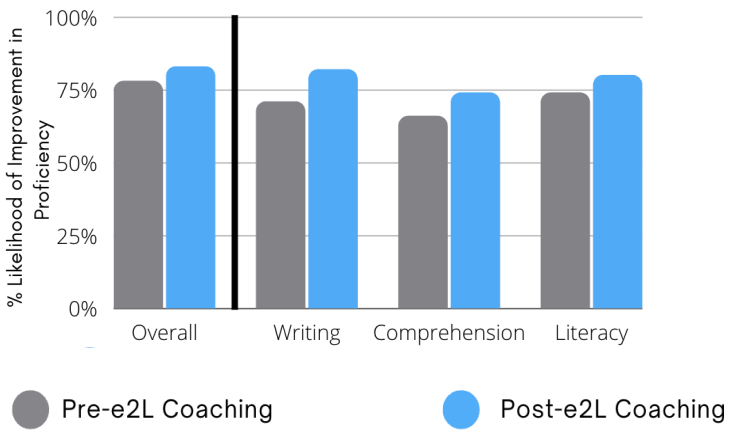
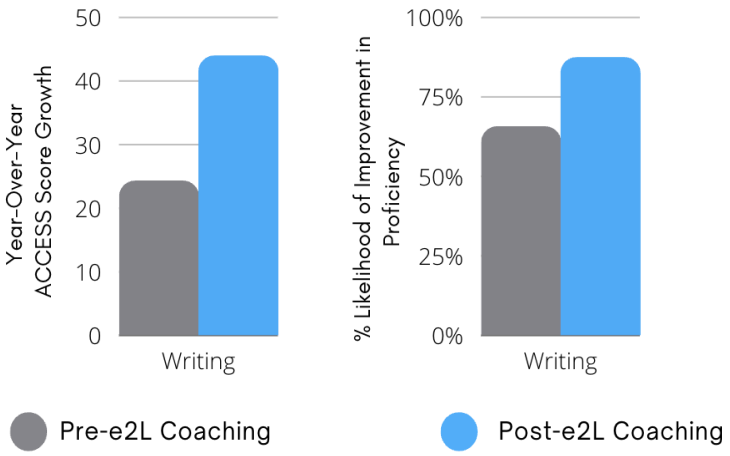


Figure 4. Third Grade Writing and Reading Improvements Pre- and Post-Coaching

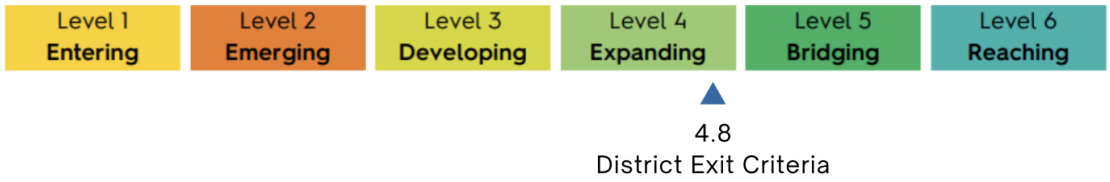
Figure 5. Fourth Grade Writing Improvements Pre- and Post-Coaching



Higher Likelihood of Growth in Proficiency for Mid-Level Students (Level 3-4)

Proficiency levels serve as interpretive scores, distinct from Scale Scores, to gauge a student's English language skills within the WIDA framework. These levels classify a student's language proficiency into one of six stages (Figure 6). A proficiency level score comprises a whole number followed by a decimal. The whole number signifies the specific proficiency level attained by the student, while the decimal indicates the extent of progress within that level. For instance, a student with a score of 3.7 is operating at proficiency level 3 and has progressed significantly within that level, being over halfway towards achieving proficiency level 4. In the district, a score of 4.8 indicates that a student can exit MLL services.

Figure 6. Six Stages of Language Proficiency on WIDA

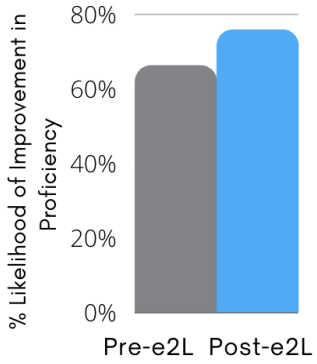


This analysis grouped students into Low, Mid, and High proficiency groups based on their baseline proficiency level according to the above scale. Low proficiency was characterized as Level 1-2, Mid proficiency includes Levels 3-4, and High proficiency includes Levels 5-6. Note that a student may be in multiple groups depending on their domain and composite test scores. Students in the Pre-e2L Coaching group who started at a low proficiency showed a similar likelihood for improvement as those in the Post-e2L Coaching group. However, students who started at a mid-level[4] proficiency

[4] t(430) = 2.21, p<.05, Cohen's d (Effect Size) = .21

demonstrated a significantly higher likelihood of overall proficiency improvement with their Post-e2L Coaching group than the Pre-e2L Coaching group (Figure 7).

Figure 7. Increased Likelihood of Improvement by Mid-Level (3-4) Students Pre- and Post-Coaching



Higher Overall Growth Year-Over-Year

Results indicate that e2L coaching had a positive impact on year-over-year gains in English language proficiency, with some variations across grade levels and domains (Tables 5-6). As a combined group, students showed significantly higher gains on WIDA ACCESS Scale Scores after teachers received e2L coaching in Overall, Speaking, and Oral Language composite scores (which includes listening and speaking components). In all other domains (Writing, Reading, Listening), and composite scores (Literacy and Comprehension), students taught by teachers post-e2L coaching also had higher year-over-year average gains, though the differences did not reach statistical significance. Additional subgroup analyses for this measure are located in the appendix.

Table 5. Student Outcomes by Year-Over-Year Gains by Domain

Instructor Coaching	Overall Score	Oral Language	Comprehension	Literacy
Pre-e2L	**41.1**	**32.8**	44.5	44.7
Post-e2L	**46.8***	**40.3***	49.4	49.6
T-Statistic	**2.24**	**2.44**	1.45	1.61
Effect Size	**.14**	**.15**	n/a	n/a

Note: *p<.05, **p<.01, ***p<.001.

Table 6. Student Outcomes by Year-Over-Year Gains, Subscales

Instructor Coaching	Speaking	Listening	Reading	Writing
Pre-e2L	**25.6**	39.9	46.6	41.8
Post-e2L	**36.6***	43.9	51.7	47.5
T-Statistic	**2.72**	1.12	1.28	1.85
Effect Size	**.17**	n/a	n/a	n/a

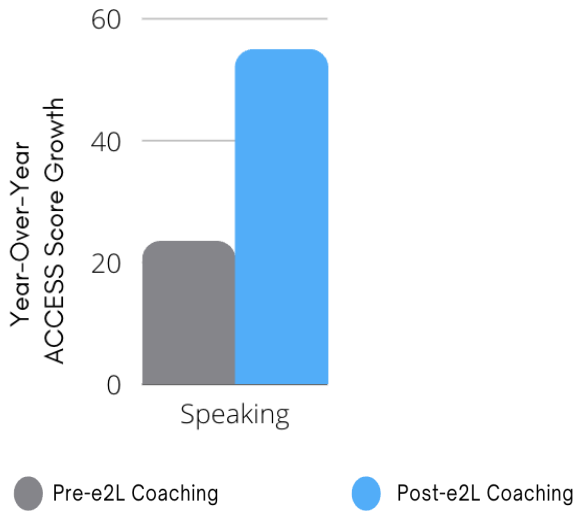
Note: *p<.05, **p<.01, ***p<.001.

Grade-Level Year-Over-Year Gains

The biggest impact on students taught by e2L-coached teachers was seen in second and third grade. First and fifth-grade students did not demonstrate significant year-over-year growth in any language domain and fourth-grade students taught by e2L-coached teachers showed a significant improvement in writing only. Additional subgroup analyses for this measure are located in the appendix.

Second Grade: While second-grade students did not show significant overall year-over-year gains, there were significant improvements in speaking and oral language, which includes speaking components.

Figure 8. Second Grade Speaking Gains Pre- and Post-Coaching



Third Grade: Teachers who received e2L coaching had a significant positive impact on third-grade students. They achieved significantly higher overall Scale Score gains compared to the previous year. There were also significant year-over-year gains in Writing (Figure 9).

Table 7. Grade-Level Significant Effects by Domain

Grade	Subtest	Statistical Result	Effect Size
2	Speaking	$t(130) = 3.75,\ p < .01^{**}$	Cohen's d = .62
3	Writing	$t(290) = 3.12,\ p < .05^{*}$	Cohen's d = .36
4	Writing	$t(138) = 3.75,\ p < .01^{**}$	Cohen's d = .63

Figure 9. Third Grade Writing Gains Pre- and Post-Coaching

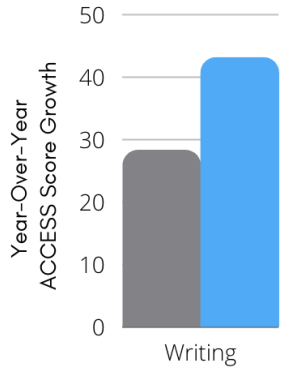
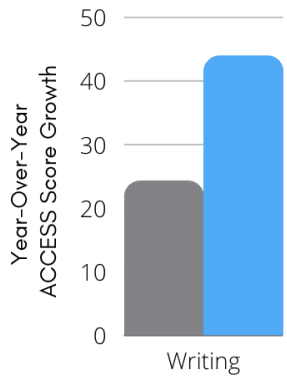


Figure 10. Fourth Grade Writing Gains Pre- and Post-Coaching



E2L Badge Attainment

After observing an overall effect of e2L-coaching vs non-coaching on student outcomes in proficiency and year-over-year Scale Score gains, we took a deeper look at the nuances of e2L coaching (i.e. number of badges, badge level, etc) on student WIDA ACCESS outcomes between 2022 and 2023. Of the 38 teachers included in the longitudinal analysis, 31 attained level-one badges only. Four teachers had at least one level-two badge, and three teachers had at least one level-three badge. This made it difficult to get a true picture of the effect of depth of badge attainment on student outcomes. In addition, when looking at the number of badges earned per teacher, what we would describe as the breadth of badge attainment, there was not a clear relationship between the number of badges and

student year-over-year growth. We looked at differences between having 1, 2, 3, 4, and 5+ badges as well as subdividing the data in terms of a combination of teacher depth and breadth badge statistics. Unfortunately, this did not yield a consistent pattern of results, potentially due to the issues described above regarding the lack of variation in badge depth (levels earned) and number of badges attained. See the Appendix for more details.

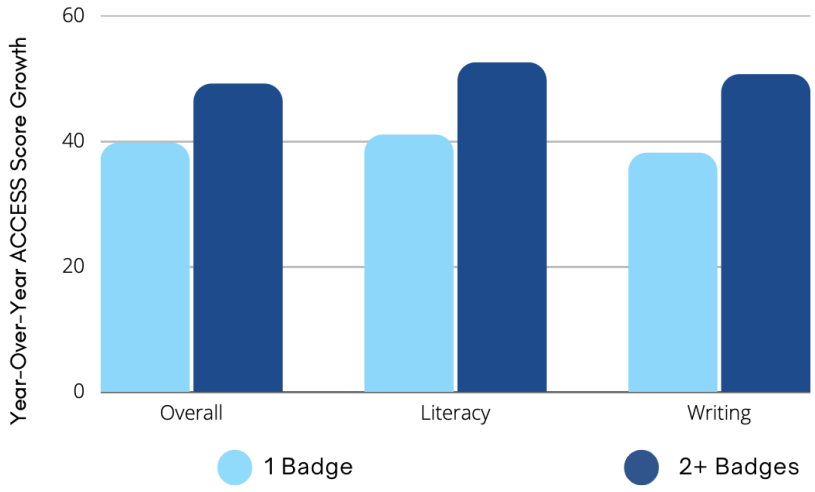
Higher Year-Over-Year Gains with Two or More Badges

We simplified the analysis to consider any differences between having one badge and having more than one badge. Note that having a single badge indicates that the teacher was still in the first level towards mastery of that standard. Students showed significantly higher year-over-year gains with teachers who had two or more badges compared to teachers who only had one badge in Overall Scale Score, Literacy, and Writing.

Table 8. Year-Over-Year Significant Effects among Coached Students by Badge Attainment

Domain/Subscale	Statistical Result	Effect Size
Overall	$t(230) = 2.36, p < .05^*$	Cohen's d = .24
Literacy	$t(233) = 2.51, p < .05^*$	Cohen's d = .26
Writing	$t(210) = 2.64, p < .05^*$	Cohen's d = .28

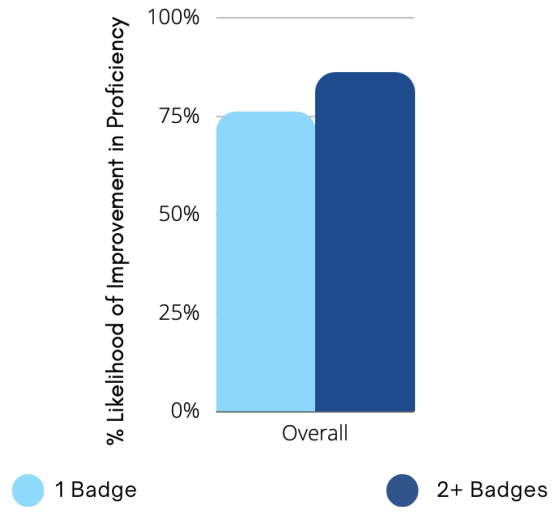
Figure 11. WIDA ACCESS Score Growth among Coached Students by Badge Attainment



Higher Likelihood of Growth in Proficiency with Two or More Badges

Students had a higher likelihood of improving their overall proficiency level with teachers who had two or more badges compared to teachers who only had one badge[5].

Figure 12. WIDA ACCESS Performance for Post-e2L Coached Students by Badge Attainment



Conclusion, Opportunities, & Implications for Future Research

This study documented e2L program implementation among MLL educators in the Providence Public School District, and impact on student outcomes. Leveraging teacher usage data and student outcome data across three years (2021, 2022, and 2023), the study provided evidence of the effectiveness of e2L coaching systems and technology tools to support the teaching of students who are learning English as a second language. The findings have value on their own to stakeholders interested in the efficacy and impact of MLL instructional materials and programs and may also serve as the foundation for future quasi-experimental research.

Assessment data in the form of Grade 1-5 WIDA ACCESS scores and proficiency levels was used in our main analyses. By comparing the change in students' WIDA Access scores year-over-year both before the educators received e2L coaching (i.e., from 2021-2022) and after (i.e., from 2022-2023), this study compared the MLL teacher's relative performance from before and after e2L Coaching. The results were generally positive -

- Engage2learn coaching was associated with a positive change in the likelihood of students to improve in overall proficiency level from grade to grade.

[5] $t(170) = 2.06, p < .05^*$, Cohen's d (Effect Size) = .23

- Likewise, students of teachers post-e2L coaching showed stronger gains on Overall Scale Scores than the students of these teachers before participating in the program.
- Educators who earned two or more e2L badges had students that showed stronger Access Scale Score growth than e2L coached teachers with only 1 badge earned.

Educators' top-earned badges were centered around the topics of Differentiation and Scaffolding; Communication; Assessment and Formative Feedback; Relevance and Authenticity; and Culture, Environment and Professional Ethics. The outcomes from the teacher survey reflect a notably positive sentiment, with all 13 respondents expressing their active integration of Engage2learn coaching strategies into their daily teaching, recognizing the alignment between these tools and their teaching practices as well as student needs, and acknowledging the considerable support these strategies offer to ELL students across various facets of their daily activities and academic advancement.

When educators built capacity in these areas mentioned above, there was a significant benefit to their students and a direct correlation to the instructional priorities suggested by Stepanek et al. (2010) that emphasize the characteristics of a high-quality ESL program. The selection of an appropriate MLL program model that was understood by all staff and implemented with fidelity, led to gains in student achievement across different domains. This differentiation of achievement could be attributed to the higher capacity of teachers to differentiate and scaffold their instruction for learners of diverse language backgrounds. Not all ELs will make gains at the same rate and within the same domain, but they will make progress. For example, second graders who are still in the "learning to read" phase of their education achieved the most gains in speaking and oral language, while 4th graders who are now reading to learn showed the highest gains in writing.

These variations can also be seen among students of the same grade level. Differentiation is a strategy that often presents challenges to teachers with some reporting that variability of academic skills, English language proficiency and background pose significant difficulties in the classroom (Gándara, 2005). With Differentiation and Scaffolding being the top area in which teachers were coached using e2L, we can see how improving educator skill in this area positively affected student achievement. Educators are also challenged in the area of assessment , with many indicating that they lack appropriate tools to adequately assess and place their English learners (Gándara, 2005). High-quality ESL instruction must include appropriate assessment and feedback in order to monitor the progress of students and place them in suitable learning environments to ensure the best possible outcomes. Badges in assessment and formative feedback were among the top three earned by educators using e2L, again directly resulting in gains among students. These overall positive results, including the significant or trending-positive findings for each of the sub-domains in the WIDA Access test, indicate the robustness of the findings, and the general efficacy of e2L coaching.

The student achievement data are relevant for diverse stakeholder groups including e2L product development and marketing, ESL policymakers, and ESL educators and administrators. Data

summarized in the report may help e2L develop and iterate to meet needs of customers as well as add to e2L's existing evidence base, both with regard to students in general, and MLL students in particular. e2L sales and marketing teams may promote the findings of the report to customers who would be interested in and/or require research and evidence of efficacy to purchase and use their coaching and talent development products and services. The results may be used in stakeholder-responsive ways such as highlighting the uses and impacts of cultural components for student engagement as well as highlighting and perhaps building on the many different use cases for e2L materials in the MLL matrix of curriculum and equity goals.

In addition to plentiful opportunities, limitations in the study do exist and may be addressed through future research and inquiry. For example, teachers in the study were not randomly assigned to participate, and thus there was no opportunity for a comparison group during the same learning time period. Likewise, teachers with high and lower levels of engagement in the program were not randomly assigned to those conditions. A randomized or quasi-experimental study that compares users with non-users during the same school year has the potential to produce research results that meet more rigorous research requirements of Tier 1 or 2 evidence-based levels. As e2L continues to expand its evidence of impact across diverse networks of educators and their students, the quantity and quality of its evidence of efficacy will continue to grow.

References

Abedi, J. (Ed.). (2007). English language proficiency assessment in the nation: Current status and future practice (pp. 3-10). Davis, CA: University of California, Davis, School of Education.

Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). Listening to teachers of English language learners: A survey of California teachers' challenges, experiences, and professional development needs.

Lee, N. (2012). District Readiness to Implement Standards-based Reform for English Language Learners a Decade after the No Child Left Behind Act. Retrieved from https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2012_04.pdf

Najarro, I. (2023, April 18). U.S. English Learners' Language-Proficiency Scores Still Below Pre-Pandemic Years. Education Week. https://www.edweek.org/teaching-learning/u-s-english-learners-language-proficiency-scores-still-below-pre-pandemic-years/2023/04#:~:text=English%2DLanguage%20Learners-,U.S.%20English%20Learners'%20Language%2DProficiency%20Scores,Still%20Below%20Pre%2DPandemic%20Years&text=Virtual%20learning%2C%20as%20a%20result,to%20new%20national%20assessment%20data.

National Center for Education Statistics (2018). National Assessment of Educational Progress (NAEP), 2017 Mathematics and Reading Assessments. U.S. Department of Education, Institute of Education Sciences. Retrieved from the Nation's Report Card. https://www.nationsreportcard.gov/reading_2017/#/nation/achievement?grade=4.

National Center for Educational Statistics (2021). Digest of educational statistics. "Table 204.20 English learner (EL) students enrolled in public elementary and secondary schools, by state: Selected years, fall 2000 through fall 2019". U.S. Department of Education. https://nces.ed.gov/programs/digest/d21/tables/dt21_204.20.asp

National Center for Education Statistics, (2019). National Teacher and Principal Survey (NTPS), "Public School Teacher Data File," 2017–18. U.S. Department of Education.

Sahakyan, N., & Cook, H. (2021). Examining English Learner Testing, Proficiency, and Growth: Before and Throughout the COVID-19 Pandemic. https://wida.wisc.edu/sites/default/files/resource/Report-Examining-English-Learner-Testing-Proficiency-Growth.pdf

Stepanek, J., Raphael, J., Autio, E., Deussen, T., & Thompson, L. (n.d.). Creating Schools That Support Success for English Language Learners. https://files.eric.ed.gov/fulltext/ED519412.pdf

Talandis, G., Jr., & Stout, M. (2015). Getting EFL students to speak: An action research approach. ELT Journal: English Language Teaching Journal, 69(1), 11-25. doi:10.1093/elt/ccu037

Van Ngo, H. (2007). Toward Quality ESL Education. TESL Canada Journal, 24(2), 1-22.

Zen, D. (2000, November 30). What is wrong with ESL programs in schools?. ERIC. https://eric.ed.gov/?id=ED482580

Appendix

Additional SubGroup Analyses

Gender Effects

The main effect of higher year-over-year WIDA ACCESS gains seems to be driven by students identified as Male. Males had higher overall year-over-year gains in Post-e2L Coaching compared to Pre-e2L Coaching, $t(467) = 2.20$, $p < .05$, Cohen's $d = .20$. Students identified as Female students had higher mean overall gains, but it was not statistically significant. Both Males and Females showed significantly higher likelihood of improving in writing proficiency.

Students who scored 4.8 or higher on WIDA ACCESS overall proficiency were considered sufficiently proficient in English for mainstream classroom instruction, and thus "exited" from MLL programs in PPSD. 3rd-grade boys were more likely to exit post-coaching, $t(66)=2.05$, $p<.05$, Cohen's d effect size = .35. 4th grade girls are driving the main effect with 1 (out of 22) exiting pre-coaching and 13 (out of 47) exiting post-coaching. $t(67)=2.89$, $P<.01$, Cohen's d effect size = .65.

Race and Ethnicity

While 92% of the students in our sample identified as Hispanic/Latino, there was a roughly even split between students who were Hispanic-White and Hispanic-NonWhite. When considered as a whole, they drove the main effects on overall year-over-year gains and likelihood of proficiency improvement, however, there were some interesting differences when race was considered individually and in combination.

Hispanic/Latino student sample size

Subgroups	Pre-e2L Coaching Cohort	Post-e2L Coaching Cohort
Hispanic-White	241	231
Hispanic-NonWhite	240	203
Total Hispanic/Latino	**481**	**434**

Year-over-year Gains: Racial-ethnic Minority students (any student who did not identify as White) had significantly higher year-over-year gains on Overall, Speaking, and Oral composite scores. Similarly, students who identified as Hispanic-NonWhite showed higher overall year-over-year gains when taught by teachers post-e2L coaching compared to pre-e2L coaching, and this effect was not seen

for Hispanic-White students. There were no significant year-over-year effects pre- and post-e2L coaching for other races individually.

Race/Ethnicity Significant Effects for Year-Over-Year Gains

Group	Subtest	Statistical Result	Effect Size
Hispanic/Latino	Overall	$t(889)= 2.15, p < .05$	Cohen's d = .14
Minority	Overall	$t(466)=2.23, p < .05$	Cohen's d = .20
Minority	Speaking	$t(468) = 2.91, p < .05$	Cohen's d = .26
Minority	Oral	$t(468) = 2.56, p < .05$	Cohen's d = .23
Hispanic-NonWhite	Overall	$t(432)=2.17, p < .05$	Cohen's d = .21

Likelihood Proficiency Improvement: All Hispanic/Latino students taught by teachers post-e2L coaching had a higher likelihood of improved proficiency overall, as well as in writing and literacy. Minority students also showed a higher likelihood of improved proficiency overall, in writing, and all composite scores. White students taught by teachers in the post-e2L-coached cohort had a higher likelihood of improving in writing , t(498)=3.06, p<.01, Cohen's d = .27, but not in any other domains. Hispanic-NonWhite students had a higher likelihood of improved proficiency overall, in writing, and in comprehension and literacy. Hispanic-White students showed this effect for writing t(455)=3.28, p<.01, Cohen's d = .30, but not in any other domain.

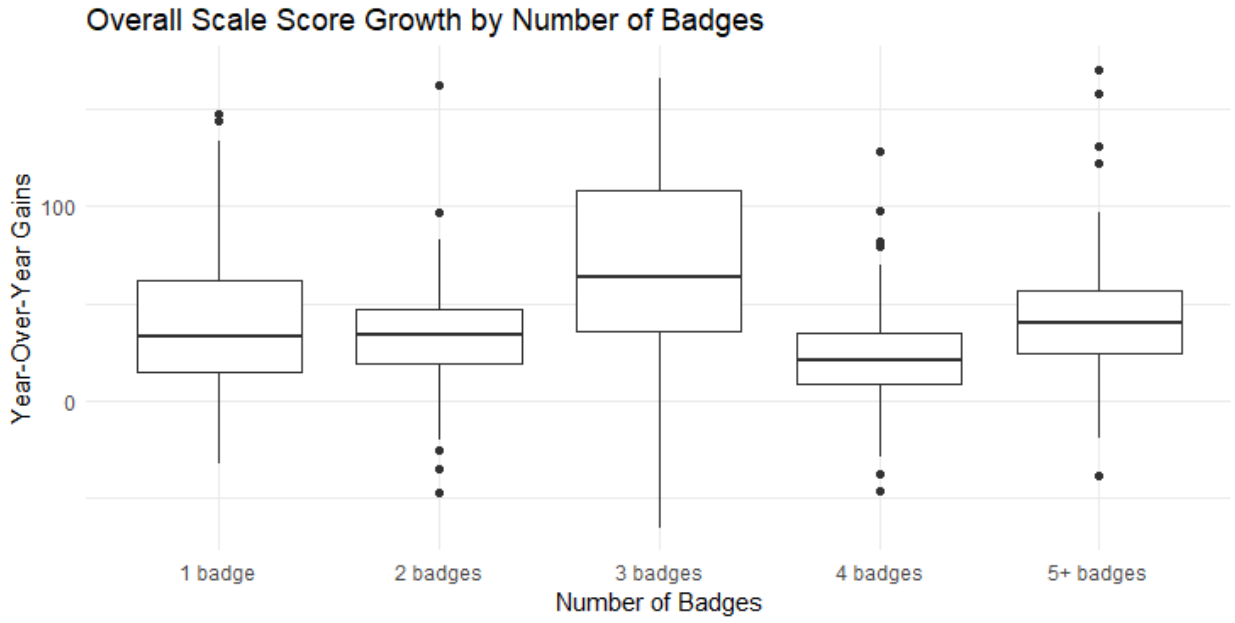
Group	Subtest	Statistical Result	Effect Size
Hispanic/Latino	Overall	$t(900) = 2.28, p < .05$	Cohen's d = .15
Hispanic/Latino	Writing	$t(896)=4.42, p < .001$	Cohen's d = .17
Hispanic/Latino	Literacy	$t(902)=2.67, p < .05$	Cohen's d = .26
Minority	Overall	$t(472)=3.15, p < .01$	Cohen's d = .29
Minority	Writing	$t(474)=3.08, p < .01$	Cohen's d = .28
Hispanic-NonWhite	Overall	$t(432) = 3.02, p < .01$	Cohen's d = .29
Hispanic-NonWhite	Writing	$t(436)=2.96, p < .05$	Cohen's d = .28
Hispanic-NonWhite	Literacy	$t(439)=2.96, p < .01$	Cohen's d = .28
Hispanic-NonWhite	Comprehension	$t(436)=2.43, p < .05$	Cohen's d = .23

Likelihood Exiting: Hispanic/Latino students taught by teachers post-e2L coaching were more likely to reach exit criteria, $t(801)=2.27$, $p<.05$, Cohen's $d = .15$. Again this effect is mostly driven by 4th grade students, $t(127)=3.08$, $p<.01$, Cohen's $d = .52$. There was no evidence of an effect of exiting for minority or White students, except for 4th grade White students. No effect for Hispanic-NonWhite on exiting, as a whole or by grade_level. The 4th grade effect is driven by Hispanic-White students, $t(69)=2.88$, $p<.01$, Cohen's $d = .66$.

Additional Exploration of Badge Attainment

When considering badges in bins (1, 2, 3, 4, and 5+), there seems to be an effect of having 3 badges. This group, consisting of 8 teachers and 137 students, was significantly different from the 1, 2, 4, and 5+ badge groups in Overall Scale Score. There was also a small difference between 4 badges and 5 or more. This effect persisted in the domain scores, Reading and Writing, as well as Listening, with the exception of 4 and 5+, and in Speaking we see significant differences between 3 and 4 badges and 4 and 5 or more badges.

Diving deeper into the three-badge effect, 50% of students taught by a teacher with 3 badges were first-grade students and 94% of the 137 students improved their overall score from 2022 to 2023.



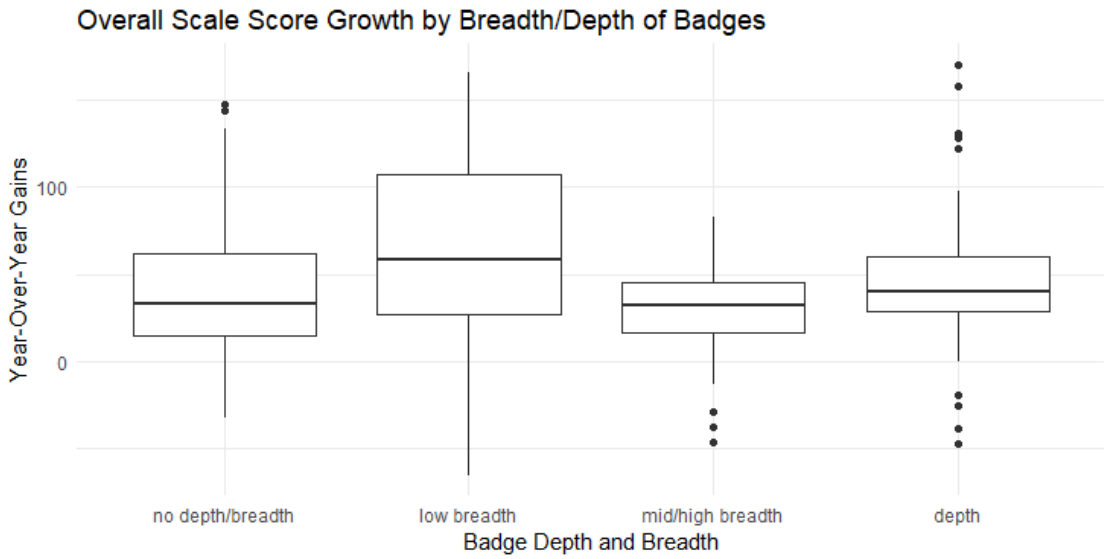
Depth vs. Breadth

We combined the number of badges and maximum badge level to try to describe teachers in terms of depth and breadth. Depth is characterized as earning higher badge levels and breadth is quantified by the number of badges earned. Given our sample of teachers, we split them in the following ways:

Group	Description	Teachers	Students
No depth or breadth	1 badge, max level 1	10	117
Low breadth only	2-3 badges, max level 1	12	151
Mid/high breadth only	4+ badges, max level 1	9	117
Depth, any breadth	Max badge level 2-3	7	84

The first three groups are all no depth (maximum badge level of 1) and differ in terms of number of badges. The final group ranges in badge number from 2 to 12 and includes any teachers who earned badges in level 2 or 3.

When grouped in this way, there was a difference between 'no depth or breadth' and 'low breadth'. In other words, just having 2 or 3 badges made a difference from just having one. This trend did not continue with having more badges as the mid/high band was actually lower (in terms of magnitude of year-over-year gains) and just 'depth' was also less than 'low breadth', but higher than 'mid/high'. This may indicate that getting more than one badge significantly increases year-over-growth, but as more and more badges are collected, the initial impact is not as high. Note that the sample sizes per group were not perfectly matched and more data may be needed to understand the effect of badge depth.


Overall Scale Score Growth by Breadth/Depth of Badges
Year-Over-Year Gains
100
0
no depth/breadth
low breadth
mid/high breadth
depth
Badge Depth and Breadth

LEARNING EXPERIENCE DESIGN | LXD Research | engage2learn

LXD Research is an independent research firm that evaluates educational programs with ESSA-aligned methods.

Learn more at www.lxdresearch.com

For additional information about **growelab** visit:

visit www.engage2learn.org

engage2learn

Learning Experience Design (LXD) Research & Consulting

A division of Charles River Media, Inc.