BEAVERTON SCHOOL DISTRICT

Bilingüismo es un superpoder:
Exploring the long-term results of Beaverton's dual language programs

Manuel Vazquez Cano and Jason Greenberg Motamedi

May 2024

Beaverton School District is home to more than 13,215 multilingual K–12 students—one-third of the entire district student population—who speak more than 110 languages. Among these linguistically and culturally diverse students, 5,438 are classified as English learners. These students have substantial assets, such as multilingualism and multiculturalism, that enrich their classrooms and communities.

Participation in Beaverton's dual language programs appears to have a positive academic impact on students, with the greatest benefits for students who were classified as English learners.

Compared to other similar students, dual language students classified as English learners:

- **Scored higher on state English language arts and math assessments**
- **Had higher rates of reclassification after grade 8**
- **Were more likely to be on track to graduate from high school**

To leverage those assets and best serve diverse multilingual and multicultural students, Beaverton School District has invested in dual language education. Beaverton launched its first dual language classroom in 1992 at Barnes Elementary School. In spring 2023, the dual language program had grown to serve 1,902 K–12 students in two elementary schools, one K–8 school, two middle schools, and three high schools.[1] These programs teach content in English and Spanish, with the goal of developing multilingual, multiliterate, and multicultural students. For more details about Beaverton School District's dual language programs, including a summary of each school's history and program models, see appendix C.

[1] Beaverton also has two dual language charter schools—one offering instruction in Mandarin Chinese and the other in Spanish. Since these schools are outside the direction of the Beaverton multilingual department, they are not included in our analysis.

Since 2015, Beaverton School District has partnered with Education Northwest to support programs for multilingual learners through research and technical assistance. One outcome of this partnership was a study (Greenberg Motamedi et al., 2019) which found that students classified as English learners who participated in dual language or co-teaching programs made greater English language proficiency growth than similar students in pull-out English language development programs or whose parents waived services.

Building off that study and other national research, Education Northwest and the Beaverton School District multilingual department, in consultation with Virginia Collier and Wayne Thomas, codesigned a rigorous longitudinal evaluation of district dual language programs. The goals of this evaluation were to:

- Create a pre-COVID-19 pandemic baseline to track the outcomes of Beaverton's dual language programs over time
- Provide Beaverton School District with data that informs its dual language program planning
- Build evidence of dual language program near- and long-term outcomes

Our research addresses the following question: *How do English learner students who participate in Beaverton's Spanish-English dual language programs perform on academic outcomes, including student assessment scores, reclassification rates, and graduation, in comparison to peers who do not participate in dual language programs?*

Who is an English learner?

Throughout this report, we use two terms to describe students' English learner status:

English learner students include all students who have ***ever*** been classified as English learners regardless of if or when they exited the program and reclassified as former English learners.

In any given school year this includes students who are currently classified as English learners, former English learner students who were reclassified as "English proficient," and students being monitored for their first few years after reclassification. In our sample, all English learner students were identified (classified) in kindergarten.

English learner students are eligible for a variety of services and supports, including English language development, although their families may waive these services.

Never English learner students includes all students who were ***never*** classified as English learners at any point in their K–12 career, including both monolingual and multilingual English speakers.

Methods

To understand the impact of participating in a dual language program, we compared academic outcomes—including Oregon English language arts, math, and English language proficiency state assessment scores; reclassification rates; being on track to graduation in grade 9; high school graduation; and postsecondary participation—of students in a Beaverton dual language program to the outcomes of similar Beaverton students who did not participate in a dual language program.

With dual language programs, it can be challenging to evaluate impact because the students and parents who choose to participate may have distinct backgrounds from their peers who do not participate. For example, parents with higher levels of education may choose to enroll their child in a dual language program because they understand the potential benefits of bilingualism and biliteracy. Parents may also choose the program because they speak a language other than English at home and want to maintain their child's connection to their language and culture.

These differences can make it difficult to distinguish whether changes in student outcomes can be attributed to participating in the dual language program or to the characteristics of participating students and parents. In Beaverton, 97 percent of dual language participants who were classified as English learners identified as Latinx[2] and spoke Spanish at home, compared to 58 percent of English learner students not in dual language programs who spoke Spanish at home. Dual language English learner students also had lower English language proficiency than students who did not participate in dual language when they entered kindergarten, as measured by the Woodcock-Muñoz assessment (see table A1 in appendix A).

Why did we use matching?

Random assignment is the most rigorous method of determining the impact of a program. Random assignment of enough participants ensures that all characteristics that could affect student performance are balanced and equal between students who participate in the program and those who do not. However, random assignment is not always practical or ethical. Since we could not randomly assign students to dual language programs, we used statistical methods to create a matched sample of students who are similar to dual language participants in key areas—except that they did not participate in the program.

To determine whether participation in Beaverton's dual language program had an impact on student outcomes, we needed to find a comparison group of students who were very similar to dual language students but did not participate in the program. To do this and reduce selection bias in this study,

[2] Throughout this report we use Latinx rather than Latina or Latino to avoid gender-specific labels.

we used statistical matching methods to identify a group of comparison students in Beaverton schools who were statistically indistinguishable from dual language students when they enrolled in kindergarten.

We started with all Beaverton kindergartners who participated in dual language programs for at least two years, then used coarsened exact matching to find comparison students who did not participate in the dual language program (Iacus et al., 2012). This statistical matching method ensured that each dual language student had the exact same demographic characteristics—race/ethnicity, gender, and home language—as their non-dual language peers, as well as special education identification and kindergarten entry year. They also had very similar initial English language proficiency levels when they entered kindergarten (i.e., differences were very small and not statistically significant). Matching students before participation in the dual language program allows us to better disentangle the potential impact of dual language from other influences, such as students' home environment and prior learning. However, without random assignment we cannot be certain that the differences we observe result from the dual language program.

For more detail on the methods used in this study, see appendix A.

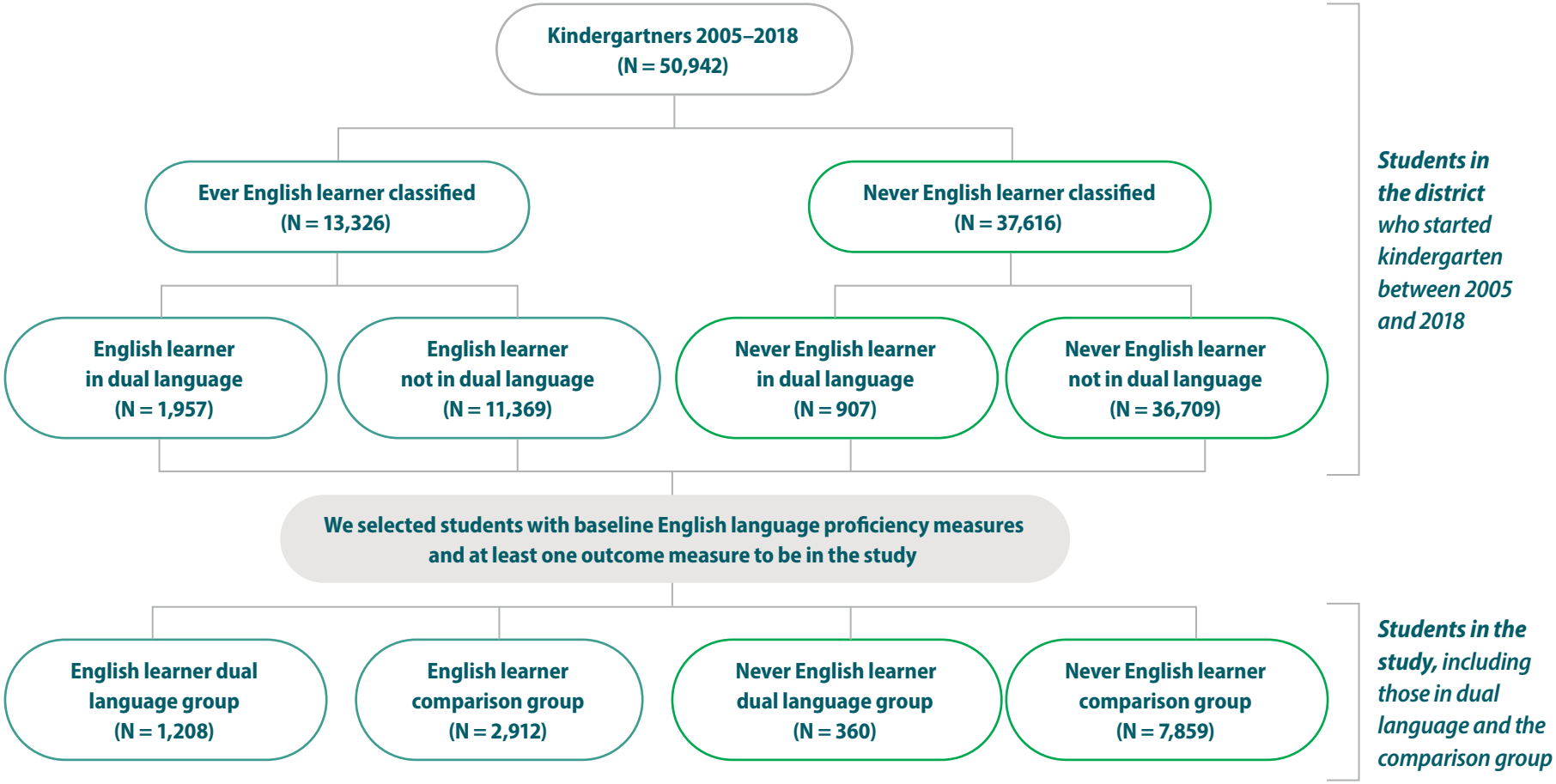
Our sample

We used student-level administrative data from Beaverton School District to track the outcomes of 50,942 students in 14 cohorts who entered a district kindergarten from fall 2005 (cohort 1) to fall 2018 (cohort 14; see tables A1 and A2 in appendix A). **We identified 1,208 students who were classified as English learners in kindergarten, enrolled in a dual language program for at least two years, and had a baseline measure of their English language proficiency.**[3] These students also had at least one of the outcome measures, such as assessment scores, reclassification date, or graduation date (figure 1).

We also identified 360 students who enrolled in a dual language program in kindergarten, participated for at least two years, had at least one of the outcome measures, and were never classified as English learners. Some of these students were multilingual and spoke English and other languages at home, and others spoke only English at home.

[3] For most students, the baseline measure of English language proficiency in listening, speaking, reading, and writing was the Woodcock-Muñoz Language Survey, typically taken within 30 days of kindergarten enrollment. However, a small number of students (N = 739) who began kindergarten in fall 2007 and 2008 were missing this score. In its place, we used their end-of-kindergarten English language proficiency assessment (in spring 2008 and 2009, this was the Oregon ELPA).

Figure 1. The study sample included students who enrolled in dual language programs from 2005–2018



Notes: The N values indicate the total number of unique students in the entire analytic dataset. Each analysis has a different number of students. See table A1 in appendix A for complete data.

Source: Education Northwest analysis of Beaverton School District administrative data, 2005–2018.

The comparison group

We also identified a group of 2,912 students who were classified as English learners in kindergarten and never enrolled in a Beaverton dual language program. These students also had a baseline English language proficiency measure and at least one of the outcomes.

To ensure an "apples-to-apples" comparison, **we used statistical methods to match each dual language student to students not enrolled in the dual language program.** Comparison group students had the exact same home language, race, and gender as the dual language students. To ensure that students started at the same point linguistically, we matched English language proficiency scores so that the dual language and comparison groups had identical scores.[4] (For more details about our matching methods, see appendix A.) There were no statistically significant differences between these groups, establishing baseline equivalency (see tables A3–A12 in appendix A).

We also selected a group of students who were never classified as English learners and participated in dual language, then matched them to a comparison group of students who were also never classified as English learners but did not participate in dual language and were otherwise indistinguishable when they entered kindergarten.

What is baseline equivalence?

We wanted to make sure that the only difference between dual language and comparison group students was that one group participated in dual language and the other did not.

Baseline equivalence is a way of measuring how similar the groups were just as they began kindergarten. It ensures that the characteristics that could affect future achievement are the same. These include eligibility for special education, race/ethnicity, and prior academic achievement.

We used effect size—a measure based on standard deviation units—to quantify the difference between groups and establish baseline equivalence.

- **Groups are equivalent when effect sizes are 0.05 or less**
- **Groups are equivalent with statistical adjustment when effect sizes are between 0.05 and 0.25**
- **Groups are not equivalent when effect sizes are greater than 0.25**

(What Works Clearinghouse, 2015)

[4] Ideally, we would be able to match students on measures of academic achievement early in their kindergarten year. Unfortunately, we did not have a measure of prior academic achievement. Instead, we used students' beginning English language proficiency scores. Beginning English language proficiency is correlated with measures of academic achievement such as math (Abedi & Lord, 2004) and so serves as a proxy measure for achievement.

Findings

Dual language students who were ever classified as English learners outperformed their peers who were not enrolled in dual language. Their assessment scores were 1.5–10.5 percentile points higher in English language arts and 2.7–13.2 percentile points higher in math. Participation in dual language also slightly increased the likelihood that students were reclassified as former English learners by the time they entered high school. By grade 9, 88 percent of dual language English learner students reclassified, 3 percentage points higher than similar students who did not enroll in dual language programs.

Dual language students who were never classified as English learners—including those from multilingual and English-only homes—performed as well as, or slightly better than, their peers who were not enrolled in dual language programs on English language arts and math assessments. However, these differences were not statistically significant, so we cannot say with confidence that participation in dual language programs caused the difference.

Finally, dual language English learner students were more likely to be on track to high school graduation at the end of grade 9. A higher percentage of dual language students graduated from high school and attended a two- or four-year college; however, again these differences were not statistically significant and we cannot be certain that participation in dual language programs caused this difference.

These findings are explored below.

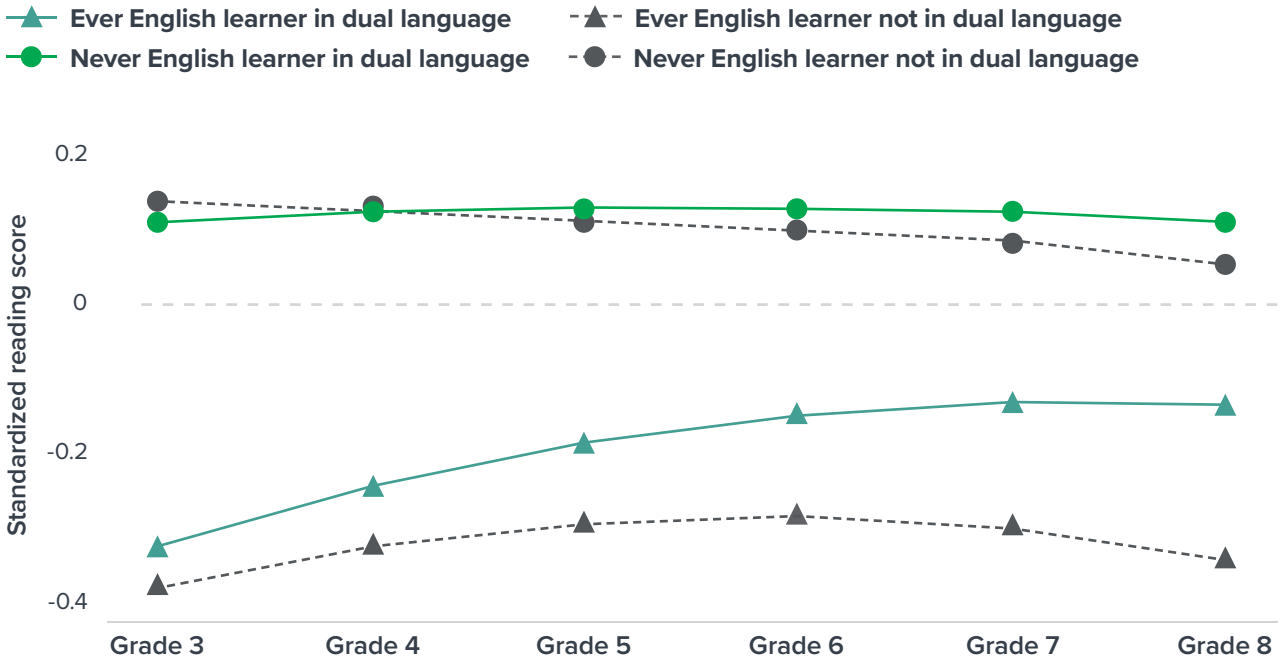
Dual language students classified as English learners scored higher on state English language arts assessments than their peers

Participation in a dual language program was positively related to English learner students' English language arts achievement. **English learner students in Beaverton dual language programs showed faster growth on the state English language arts assessment than their peers who did not participate in dual language programs** (figure 2). By grade 6, the beginning of middle school, English learners in dual language programs (shown in figure 2 as teal triangles) had grown significantly faster, by 0.17 standard deviations, than their peers who did not participate in dual language programs (grey triangles).

What is a standard deviation?

A standard deviation is a way of expressing how far a score is from the mean, or in this case how much participating in dual language moved student scores from where they would be if students had not participated in dual language.

Figure 2. Dual language program participation increased English language arts growth among English learner students



Note: The figure represents the model-predicted standardized reading scores.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2016. See table B2 in appendix B for complete data.

This growth continued through high school (see table B2 in appendix B), and **participation in a dual language program appeared to narrow the opportunity gap between English learner students and their peers who were never classified as English learners.** In grade 8, the gap between English learners and never English learners was narrowed by about half—from 0.40 standard deviations among students not in dual language to 0.21 standard deviations among students in dual language.

Dual language participants who were never English learners (shown in figure 2 as green circles) also had faster growth than their peers who did not participate in dual language programs (grey circles). However, the difference between students never classified as English learners was small. In other words, participating in dual language appears to be related to positive growth in English language arts scores regardless of students' English learner status, but growth was much greater among students ever classified as English learners.

English learner students enrolled in a Beaverton dual language program scored higher on the state English language arts assessment than their peers who did not participate. In grades 6, 7, and 11 these differences were statistically significant, suggesting that participation in a dual language program was correlated to the higher score (figure 3). For example, English learner dual language students had higher Oregon English language arts assessment scores than their peers not in dual language, by 9.7 percentile points in grade 6, the equivalent of moving a student in the 50th percentile to the 59.7th percentile.

There were differences between the scores of students enrolled in a dual language program and their peers in grades 3, 4, and 8, but they were not statistically significant. This means that we cannot be confident that differences between these groups exist. Similarly, differences in grade 5 were "promising" ($p \leq 0.10$) but did not reach the level of confidence (at least 95%, or $p \leq 0.05$) for us to be sure the differences are meaningful.

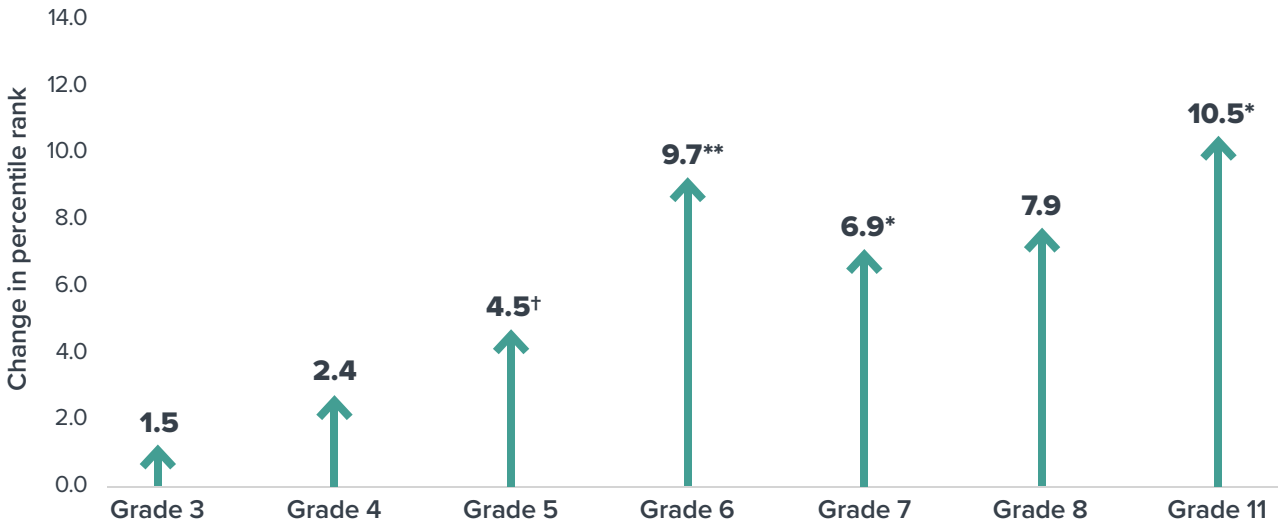
What is statistical significance?

We need to be confident that the differences we see between dual language students and their matched peers are meaningful and related to participation in dual language programs. We want to be sure that the differences are not due to chance, normal variation between students, or other differences between the groups. To determine this confidence, we use a *p*-value. We are reasonably confident when a *p*-value is equal to or less than 0.05 (in other words, we are 95% confident).

Sometimes we find that differences are visible but not significant. This means we cannot be as confident as we would like that program participation was related to the difference. This does not mean program participation was not related to the outcome; we just cannot say the two were related with a high level of confidence.

In the figure on the next page, we use one or more asterisk (*) to indicate statistical significance ($p < 0.05$), and a dagger symbol (†) to indicate marginal or promising significance ($p < 0.10$).

Figure 3. Dual language program participation increased English language arts scores among English learner students in middle and high school



† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Note: The figure represents the model-predicted standardized English language arts score transformed into an improvement index, as described by the What Works Clearinghouse (2022).

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2016. See table B3 in appendix B for complete data.

To contextualize these effects, we identified two interventions in the Institute of Education Sciences' What Works Clearinghouse that had positive impact on English language arts scores for English learner students in similar grades. Both interventions required large investments in teacher professional development, and both appeared to have smaller impact on student assessment scores than participation in Beaverton's dual language programs.

- Pathway to Academic Success trains teachers to incorporate cognitive strategies into reading and writing instruction for English learner students in grades 6–12. The program includes 46 hours of training for all participating teachers as well as coaching. *Pathway to Academic Success had a smaller impact on overall English language arts scores than participation in dual language,* with a 3 percentile point improvement to middle and high school students compared to the 6.9–10.5 percentile point improvement we found in grades 6, 7, and 11 (Kim et al., 2011).
- The New Teacher Center's induction model provides new teachers with two years of coaching and mentoring for at least three hours per month. *The New Teacher Center intervention had a smaller impact on overall English language arts scores than participation in dual language,* with a 4 percentile point improvement for a large group of that included students classified as English learners (Young et al., 2017).

Dual language students classified as English learners scored higher on state math assessments than their peers

Participation in a dual language program was positively related to students' math achievement. Students classified as English learners who enrolled in a Beaverton dual language program showed faster growth on the state math assessment than their peers who did not participate in dual language programs.

As with English language arts, dual language English learner students grew faster in math, by more than 0.17 standard deviations in grade 6, than their peers who did not participate in dual language programs (figure 4).

Faster growth in math among dual language students continued in high school and appeared to narrow the opportunity gap between English learner students and their peers who were never classified as English learners (see table B4 in appendix B). In grade 6, the opportunity gap between English learners and never English learners in math was narrowed by about half, from 0.42 standard deviations among students not in dual language to 0.24 standard deviations among students in dual language.

Dual language students who were never classified as English learners also outperformed their peers who did not participate in dual language programs, but not by the same margins.

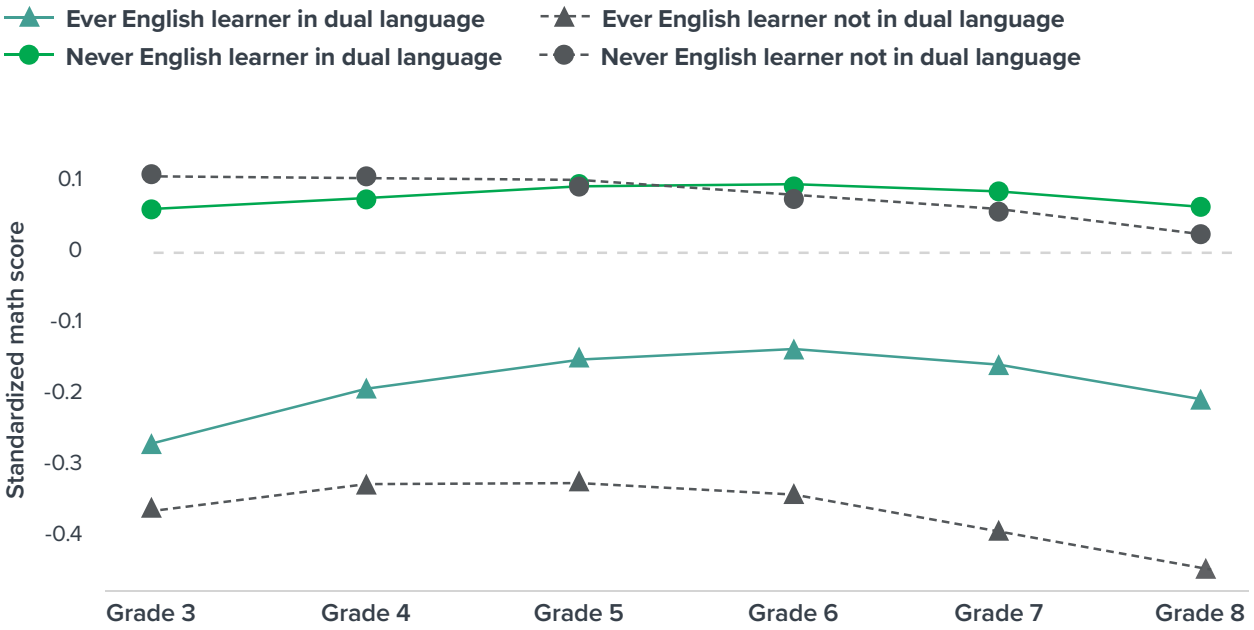
How do we know that dual language programs narrow the opportunity gap between students who were ever classified as English learners and those who were never classified as English learners?

Despite their academic, linguistic, and cultural strengths, English learner students perform less well—on average—than never English learners on many educational outcomes and assessments (Sugarman & Geary, 2018). We believe that these differences are driven not by ability but by limited access to equitable and high-quality educational content. We also believe that dual language programs are more likely to provide students with equitable, high-quality educational content.

Figures 2 and 4 illustrate this difference, with English learner students performing less well than never English learners. However, the gap between English learners and never English learners in dual language programs is about half the size of the gap between the same groups of students who did not participate in dual language programs.

Dual language programs not only improve English language arts and math outcomes for all students but also act as a tool for equity, diminishing the opportunity gap without affecting the performance of students never classified as English learners.

Figure 4. Dual language program participation increased math growth among English learner students

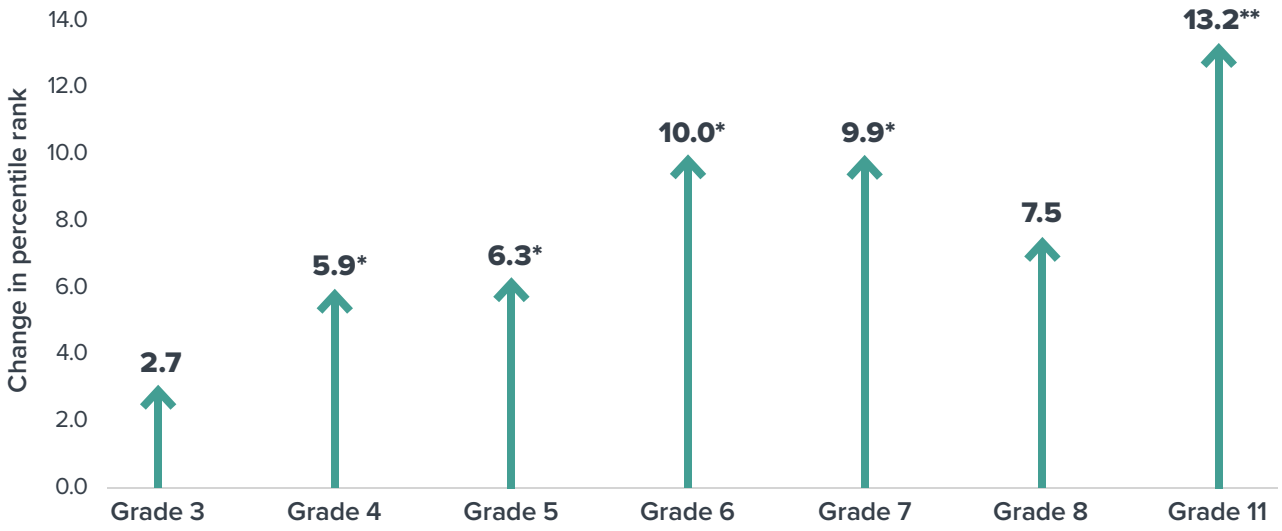


Note: The figure represents the model-predicted standardized math scores.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2016. See table B4 in appendix B for complete data.

Students ever classified as English learners who enrolled in a Beaverton dual language program scored higher on the state math assessment than their peers who did not participate in dual language programs. In grades 4–7 and 11, these differences were statistically significant, suggesting that participation in dual language programs was related to these improvements (figure 5). For example, dual language English learner students had higher scores on the Oregon math assessment than their peers not in dual language by 5.9 percentile points in grade 4 and 13.2 percentile points in grade 11. Similar to English language arts scores, the differences in grades 3 and 8 were not statistically significant, so we cannot be sure that the differences between groups are meaningful.

Figure 5. Dual language program participation increased math scores among English learner students at all grade levels



* $p < 0.05$, ** $p < 0.01$

Note: The figure represents the model-predicted standardized math score transformed into an improvement index, as described by the What Works Clearinghouse (2022).

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2016. See table B5 in appendix B for complete data.

We identified two interventions in the What Works Clearinghouse that had positive impact on English learners' math assessment scores in similar grades. Both interventions showed smaller impact than dual language participation in grades 6, 7, and 11.

- The New Teacher Center's induction model, described above, had a 6 percentile point impact on math assessment scores, which was about *the same impact as dual language in grades 4 and 5 but less impact in grades 6, 7, and 11* (Young et al., 2017).
- ST Math is a game-like computer program for grades 2–5 designed to minimize the need to use language in math instruction. *ST Math led to a 6 percentile point improvement in math scores for a sample of predominantly English learner students, about the same impact as dual language in grades 4 and 5* (Schenke et al., 2014).

Dual language students reclassified as former English learners at higher rates than their peers after grade 8

Participation in a dual language program was positively related to students' reclassification. **Students classified as English learners enrolled in a Beaverton dual language program were reclassified at lower rates than their peers in elementary school, caught up with their peers in grade 6, and passed them by grade 8.**

We used a statistical model (discrete time-hazard analysis; Singer & Willett, 2003) to estimate the cumulative percentage of English learners expected to reclassify as former English learners each year. After accounting for differences among students, including their initial English proficiency and demographic differences, we compared the percentage of dual language students who reclassified at the start of each grade to the percentage of non-dual language students who reclassified (see appendix A for details on this analysis).

Overall, **dual language students reclassified at lower rates in elementary grades compared to similar students not in dual language programs** (figure 6). For example, 28 percent of non-dual language English learner students were expected to reclassify by the beginning of grade 3, compared to 20 percent of dual language students (see table B6 in appendix B). In other words, K–5 students who did not participate in dual language programs were more likely to be reclassified than dual language students. This "lag" in the English language development of dual language students has been described in in other research (Steele et al., 2017; Umansky & Reardon, 2014) and is likely related to the process of language development in a dual language program (Howard et al., 2018) as well as the limitations of monolingual assessments (Sanchez et al., 2013).

Why is reclassification important?

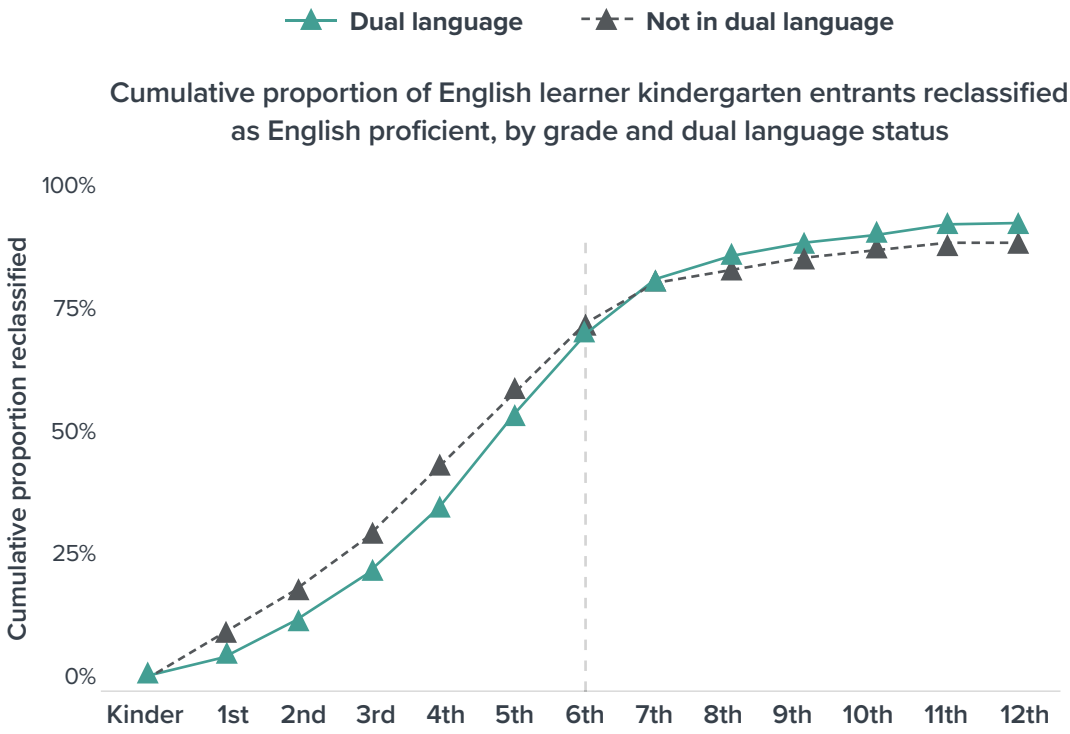
All students classified as English learners are assessed annually until they reach a grade-specific English language proficiency score in speaking, reading, writing, and listening. When they reach those scores, students are reclassified as former English learners and are no longer eligible for English learner services. At this point, they are monitored for three years to ensure that they are successful in mainstream classrooms.

Reclassification before high school entry is important because it opens up students' schedules, allowing them to take required or elective courses instead of English language development coursework (Estrada, 2014). This increases the likelihood of English learners being on track to high school graduation (Johnson, 2019) and graduating on time (Carlson & Knowles, 2016).

We found no meaningful differences in the cumulative reclassification rate during middle school between English learners who participated in dual language programs and those who did not, with about 80 percent of all students expected to be reclassified by the beginning of grade 7 (see table B6 in appendix B).

The cumulative reclassification rate of dual language students passed the rate of their peers who were not in dual language programs in grade 8, just before high school entry, and appeared to maintain its higher trajectory throughout high school. At high school entry in grade 9, 85 percent of non-dual language students had reclassified (shown in figure 6 as grey triangles) compared to 88 percent of dual language students (teal triangles). While this difference may appear small, these data do show that dual language programs correlate to only positive long-term impacts on English language development.

Figure 6. Dual language participants were more likely to be reclassified as former English learners by high school than their peers who were not in dual language programs



Note: The figure represents the model-predicted cumulative reclassification rates for English learner classified students.

Source: Education Northwest analysis of Beaverton School District administrative data, 2005–2019. See table B6 in appendix B for complete data.

A higher percentage of dual language students ever classified as English learners were on track to graduate from high school than their peers

Participation in a dual language program has a promising relationship to graduation outcomes. **Compared to their peers who did not participate in dual language programs, a higher percentage of dual language students ever classified as English learners were on track to graduate high school in grade 9 (77% compared to 71%), graduated high school (84% compared to 81%), and enrolled in a two- or four-year college (59% compared to 53%;** figure 7). However, after accounting for demographic characteristics, initial English proficiency, and other important characteristics, only the on track to graduation measure reached the level of statistical significance ($p \geq 0.05$) at which we can confidently say that there is a meaningful difference between English learner students in dual language compared to those not in dual language.

Why does the number of students matter when measuring statistical significance?

The number of students in an analysis determines the sample's accuracy, or how different the sample may be from the whole population.

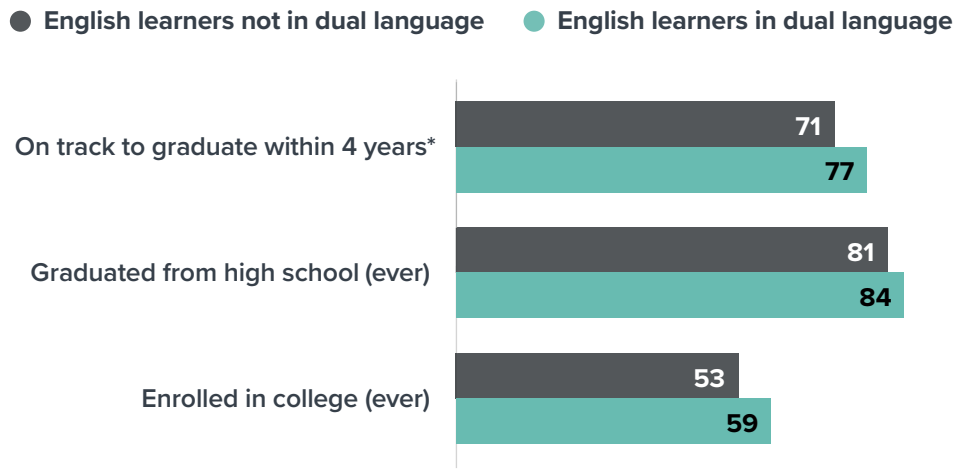
In this case, out of full sample of 13,326 Beaverton English learner students, only 277 students could have graduated high school and attended a postsecondary program by 2016.

The small sample of students means we need large changes to confidently say that dual language program participation caused this difference.

Dual language students appear to be more prepared for postsecondary success than non-dual language students; however, we cannot say with certainty that dual language program participation is the reason for this difference. It is very likely that the sample of students we could track from kindergarten to high school and beyond was too small to discern significance.

In future years we hope to bolster our analysis with additional cohorts of students to deepen our understanding of the relationship between dual language programs and postsecondary readiness and participation. We can then compare the impacts to those of other English learner-specific interventions that had impact on high school graduation or postsecondary outcomes.

Figure 7. A higher percentage of dual language students ever classified as English learners were on track to graduate, graduated, and enrolled in a two- or four-year college than their peers



* $p < 0.05$

Note: The figure represents the model predicted probabilities for English learner classified students.

Source: Education Northwest analysis of Beaverton School District administrative data including National Student Clearinghouse data, 2007–2019. See table B7 in appendix B for complete data.

Considering the results

Our evaluation found that **participation in Beaverton's dual language program correlated to growth in English learner students' English language arts and math achievement, decreased the number of long-term English learners, and had a promising relationship to graduation rates and post-secondary participation.** These findings come as no surprise: A growing body of research in Beaverton (Greenberg Motamedi et al., 2019), Oregon (Steele et al., 2017), and the nation (Morales, 2024; Umansky & Reardon, 2014) suggests that dual language programs are highly effective for students classified as English learners. When implemented well, dual language may be the most effective program option (Collier & Thomas, 2004).

Why dual language programs may benefit English learners

Dual language programs offer all students the opportunity to become bilingual, biliterate, and bicultural—one of the programs' greatest impacts, and one not measured in this evaluation. However, our evaluation suggests that dual language programs may have greater impact on English learner students, narrowing the opportunity gap between students who were ever classified as English learners and those who were not. We need more research to understand why dual language programs seem to have such powerful effects for English learners; however, existing research points to several possible explanations.

Dual language programs provide English learner students with immediate access to core content in their home language. Dual language classrooms are typically led by teachers trained to implement effective practices that support core content access (Howard et al., 2018). By contrast, students in English-only programs may have less immediate access to core content instruction since the programs focus on English language development. Increased access to core content, particularly in the early elementary years, supports academic achievement (Swanson et al., 2022).

Dual language programs provide supportive space for English learner students and language learning. Dual language programs may give English learner students greater access to English-speaking peers—and vice versa. This creates an ecosystem of mutual language development which elevates language learning (Block & Viadurre, 2019) and may counteract deficit perspectives that educators may have unconsciously internalized about English learner students (Dabach, 2014; Lee & Soland, 2023; Umansky & Dumont, 2021).

Considerations for districts and schools

Compared to other similar students, dual language students classified as English learners scored higher on state content assessments, had higher rates of reclassification after grade 8, and were more likely to be on track to graduate from high school. The impacts that dual language participation had on English language arts and math assessments, and presumably learning, appear to be larger than impacts of interventions that involved significant amounts of teacher professional development. Thus, dual language programs may be an important and worthwhile investment to improve student outcomes and make schools more equitable for English learners.

The potential benefits of dual language programs may be significant, but positive impact is not automatic. For dual language programs to be effective, they must be implemented well (Howard et al., 2018). In Beaverton School District, the impact of dual language programs results from strong leadership at the district, building, teacher, and community levels.

To build effective dual language programs, at a minimum school and program leadership must advocate for the program, adopt a specific model, and ensure its implementation; educators must be linguistically and culturally competent in both languages; and curriculum, assessments, and instruction must be linguistically appropriate and aligned in both languages (What Works Clearinghouse, 2022).

Building our knowledge for student success

Despite our growing understanding of the many social, cognitive, and economic impacts of bilingualism, we still don't know very much about what works in K–12 dual language programs (What Works Clearinghouse, 2022). For example, we don't know much about the different models of dual language education or their impacts on different populations of students, especially students who are classified as English learners or whose families are new to the United States.

For dual language programs to live up to their promise and our expectations of achievement and equity, we must increase our understanding of how and why they work. This study is one example of such an effort. By collecting and analyzing local data, districts contribute to our collective knowledge so we can implement stronger dual language programs that provide each student with access to opportunity and resources for success.

References

Abedi, J., & Lord, C. (2004). The language factor in mathematics tests. *Applied Measurement in Education, 14, 219–234.*

Block, N., & Vidaurre, L. (2019). Comparing attitudes of first-grade dual language immersion versus mainstream English students. *Bilingual Research Journal, 42*(2), 129–149.

Carlson, D., & Knowles, J. E. (2016). The effect of English language learner reclassification on student ACT scores, high school graduation, and postsecondary enrollment: Regression discontinuity evidence from Wisconsin. *Journal of Policy Analysis and Management, 35*(3), 559–586.

Collier, V. P., & Thomas, W. P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice, 2*(1), 1–20. https://www.thomasandcollier.com/s/2004-winter_njrp_astounding-effectiveness-of-dl.pdf

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social boundaries in teachers' accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk, 19*(2), 98–124.

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review, 46,* 535–573.

Greenberg Motamedi, J., Vazquez, M., Gandhi, E. V., & Holmgren, M. (2019). *Beaverton School District English language development minutes, models, and outcomes*. Education Northwest. https://educationnorthwest.org/resources/english-language-development-minutes-models-and-outcomes-beaverton-school-district

Howard, E., Lindholm-Leary, K., Rogers, D., Olague, N., Medina, J., Kennedy, B., Sugarman, J., & Christian, D. (2018). *Guiding principles for dual language education* (3rd ed.). Center for Applied Linguistics. https://www.dlenm.org/wp-content/uploads/2019/11/Guiding-Principles-for-Dual-Language-Education-3rd-edition.pdf

Iacus, S., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1–24.

Johnson, A. (2019). The effects of English learner classification on high school graduation and college attendance. *AERA Open, 5*(2). https://files.eric.ed.gov/fulltext/EJ1220740.pdf

Kim, J. S., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., Collins, P. & Land, R. E. (2011). A randomized experiment of a cognitive strategies approach to text-based analytical writing for mainstreamed Latino English language learners in grades 6 to 12. *Journal of Research on Educational Effectiveness, 4*(3), 231–263.

Lee, M. G., & Soland, J. G. (2023). Does reclassification change how English learners feel about school and themselves? Evidence from a regression discontinuity design. *Educational Evaluation and Policy Analysis, 45*(1), 27–51.

Morales, C. (2024). Dual language immersion programs and student achievement in early elementary grades. *Educational Evaluation and Policy Analysis,* https://doi.org/10.3102/01623737241228829

Sanchez, S. V., Rodriguez, B. J., Soto-Huerta, M. E., Villarreal, F. C., Guerra, N. S., & Flores, B. B. (2013). A case for multidimensional bilingual assessment. *Language Assessment Quarterly, 10*(2), 160–177. https://doi.org/10.1080/15434303.2013.769544

Schenke, K., Rutherford, T., & Farkas, G. (2014). Alignment of game design features and state mathematics standards: Do results reflect intentions?. *Computers & Education, 76,* 215–224.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford University Press.

Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion programs on student achievement: Evidence from lottery data. *American Educational Research Journal, 54*(1), 282S–306S. https://files.eric.ed.gov/fulltext/ED577026.pdf

Sugarman, J., & Geary, C. (2018). *English learners in select states: Demographics, outcomes, and state accountability policies. Fact sheet.* Migration Policy Institute. https://www.migrationpolicy.org/research/english-learners-demographics-outcomes-state-accountability-policies

Swanson, H. L., Arizmendi, G. D., & Li, J.-T. (2022). What mediates the relationship between growth in math problem-solving and working memory in English language learners? *Journal of Educational Psychology, 114*(7), 1608–1632. https://doi.org/10.1037/edu0000718

Umansky, I. M., & Dumont, H. (2021). English learner labeling: How English learner classification in kindergarten shapes teacher perceptions of student skills and the moderating role of bilingual instructional settings. *American Educational Research Journal.* Advance online publication.

Umansky, I. M., & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal, 51*(5), 879–912.

What Works Clearinghouse. (2015). *WWC standards brief.* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_baseline_080715.pdf

What Works Clearinghouse. (2022). *Dual language programs intervention report | English language arts topic area.* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/WWC_DLP_IR-Report.pdf

Young, V. M., Schmidt, R. A., Wang, H., Cassidy, L., & Laguarda, K. (2017). *A comprehensive model of teacher induction: Implementation and impact on teachers and students evaluation of the New Teacher Center's i3 validation grant, final report.* SRI International. https://www.sri.com/wp-content/uploads/2021/12/NTC-i3-Validation-Comprehensive-Report-with-App_Final.pdf

Appendix A. Methods

Data and sample

This study used administrative data from the Beaverton School District and included 14 cohorts of students who entered kindergarten from 2005–06 to the 2018–19 school years, followed longitudinally as far as postsecondary. In addition to administrative data from Beaverton School District, we used National Student Clearinghouse data for postsecondary enrollment outcomes. These data were shared by Beaverton School District and merged into our analytic dataset using a student identification number.

After collecting the data, we undertook a thorough data cleaning process to ensure that the analytical dataset had high-quality data. During this process, we checked for duplicated observations of individual students and removed them based on specific decision rules. For instance, if a student had two assessment scores in the same year, we retained the highest score and removed the other. Additionally, we analyzed the distribution of essential variables, like students' assessment scores, to test for normalcy and score ranges. If a student's score was outside the assessment range, we excluded them from the dataset. Moreover, we examined the distribution of crucial demographic variables, such as student home language and race/ethnicity, and compared them to known public records. This process helped build confidence that our dataset was representative of the students and their outcomes.

The study focused on students who were classified as English learners in kindergarten, grouped into two categories: those who participated in a Spanish-English dual language program and those who did not. We also examined outcomes for students who were never classified as English learners, comparing outcomes of those who participated in a Spanish-English dual language program to those who did not. Table A1 summarizes descriptive information of the entire sample of students in the dataset. The final analytic samples included in each analysis vary by outcome. The procedure to select students and who is included in each analytic sample is described in greater detail in the analytic approach section.

Table A1. Descriptive statistics of the entire sample

		Ever English learners		Never English learners	
	All students	**Dual language**	**Not dual language**	**Dual language**	**Not dual language**
Number of students	50,942	1,957	11,369	907	36,709
Male (percent)	51%	48%	53%	48%	51%
Race/Ethnicity (percent)					
American Indian/Alaska Native	0%	0%	0%	1%	1%
Asian	15%	1%	24%	3%	13%
Black	3%	1%	4%	4%	3%
Latinx	24%	97%	58%	33%	10%
Multiracial	6%	0%	1%	8%	8%
Native Hawaiian/Pacific Islander	0%	0%	1%	0%	0%
White	50%	2%	12%	50%	65%
Unknown race	0%	0%	0%	1%	0%
Program participation (percent)					
Special education (in kindergarten)	8%	9%	9%	3%	8%
English learner classified	26%				
English language proficiency					
Woodcock RPI (0-100)	62.01	44.48	51.69	97.37	97.66
Home language (percent)					
English language	67%	1%	2%	90%	91%
Spanish language	18%	97%	58%	8%	1%
Chinese language	2%	0%	4%	0%	1%
Vietnamese language	1%	0%	4%	0%	0%
Korean language	1%	0%	3%	0%	1%
Other language	11%	2%	29%	1%	6%

		Ever English learners		Never English learners	
	All students	Dual language	Not dual language	Dual language	Not dual language
Years in dual language (mean)	N/A	5.20	N/A	4.63	N/A

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A2. Descriptive statistics (numbers) of the entire sample, by cohort

			Ever English learners		Never English learners	
Cohort	Kindergarten entry year (fall)	All students	Dual language	Not dual language	Dual language	Not dual language
1	2005	3,564	133	800	54	2,577
2	2006	3,673	120	856	45	2,652
3	2007	3,573	145	896	54	2,478
4	2008	3,540	137	883	57	2,463
5	2009	3,705	179	832	67	2,627
6	2010	3,637	173	847	71	2,546
7	2011	3,721	165	832	81	2,643
8	2012	3,735	170	749	79	2,737
9	2013	3,723	138	789	61	2,735
10	2014	3,690	145	767	64	2,714
11	2015	3,640	139	803	70	2,628
12	2016	3,720	108	831	69	2,712
13	2017	3,471	115	773	68	2,515
14	2018	3,550	90	711	67	2,682

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Outcomes

This study examined six student outcomes, measured at different points in a student's K–12 and postsecondary trajectory. Math and English language arts achievement were measured using the state's standardized assessment, which is administered annually in the spring. Beginning in the 2014–15 school year, Oregon shifted from the Oregon Assessment of Knowledge and Skills (OAKS) to the Smarter Balanced assessment. To make assessments more comparable across grades and years, scale scores were standardized by year and grade to have a mean of zero and a standard deviation of one.

On track to graduation is a binary variable that equals one if a student earned one-quarter of the credits needed to graduate high school by the end of grade 9. High school graduation is a binary variable that equals one if a student ever earned a high school diploma. College enrollment is a binary variable that equals one if a student ever had any records of enrolling in a two- or four-year college anywhere in the country.

Lastly, we examined reclassification as an outcome in all grades. Reclassification from English learner services happens when students score above a specific threshold in their annual English language proficiency assessment. Currently, Oregon uses the ELPA21 to assess students annually. Based on this assessment, districts collect data to report to the Oregon Department of Education that describe if a student continues to be eligible for English learner services or will be shifted to monitoring status. We used this categorization to identify students who move from English learner classified to English learner exited status.

Analytic approach

To examine the impact of dual language program participation on academic, high school graduation, and postsecondary outcomes, the study used a quasi-experimental design using coarsened exact matching (CEM) to identify a comparison group. Students who participated in a dual language program for at least two years were matched in kindergarten or grade 1 with other students who never participated in a dual language program, then followed longitudinally.[5] To create a matched sample, we used Stata 17 CEM package to match students using exact matching on the following variables in kindergarten: race/ethnicity, special education identification, gender, reported home language, and entry year kindergarten cohort. We also matched using the students' initial English language proficiency—the Woodcock-Muñoz screener, which has a relative proficiency index that ranges from 0 to 100. This variable was coarsened into deciles. Only students with full data were considered in the matching approach.

[5] The majority of our sample (86% of observations) was observed in kindergarten. However, for some students, we did not observe them in the data until grade 1 but had records of their kindergarten English language proficiency screener scores. These students were retained in the cohort and assigned a kindergarten cohort entry equivalent to the prior year. For example, if we first observed a student in grade 1 in the school year 2011–12, then we assigned a kindergarten cohort entry year of 2010–11.

We made adaptations to the matching approach for the analysis that examined high school graduation and postsecondary outcomes. Students in the earliest kindergarten cohorts were the only students we were able to observe longitudinally from kindergarten to postsecondary. However, Woodcock-Muñoz screener scores were not consistently available for these groups. Instead of matching on the screener, we matched on the result of their annual English language proficiency assessment (ELPA assessment) administered at the end of their kindergarten school year. Ideally, matching variables are measured prior to receiving any treatment. However, we consider the kindergarten annual assessment score as a proxy for initial English language proficiency. The Pearson correlation measure between the Woodcock-Muñoz screener and ELPA composite score in kindergarten is 0.57.

For each outcome observed in each grade level, we checked for baseline equivalence between English learner students who participated in a dual language program and those who did not. We did this because the analytic sample differed slightly for each outcome. To calculate the mean difference between groups, we used a linear regression framework with analytic weights. This is represented by the following equation:

$$Cov_i = \beta_0 + \beta_1 DLI_i + e_i$$

Where Cov_i represents a baseline covariate of student *i* and DLI is a binary indicator that equals one if a student participated in a dual language program. The regression also includes analytic weights produced by the CEM matching procedure. β_1 represents the adjusted mean difference between dual language students and non-dual language students. We then calculated the effect size of the difference for each baseline measure observed. For binary indicators we calculated Cox's index. For continuous measures we calculated the Hedge's *g*. According to the What Works Clearinghouse (2015), groups are equivalent when the effect size is smaller than 0.05. If effect sizes are larger than 0.05 and smaller than 0.25, then groups can still be equivalent if statistical adjustments are made by including the baseline measure in the analytic model. Tables A3 to A12 present the baseline equivalency for each outcome and grade after matching. In all our analytic samples, baseline equivalence was met.

Table A3. Baseline equivalency growth analysis sample

	Intervention (N = 1,161)		Control (N = 9,686)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Grade level								
Has grade 3	100%	0.00	100%	0.00	0.00	0.00		
Has grade 4	84%	0.35	84%	0.37	0.00	0.37	0.00	0.00
Has grade 5	66%	0.46	66%	0.47	0.00	0.47	0.00	0.00
Has grade 6	52%	0.50	52%	0.50	0.00	0.50	0.00	0.00
Has grade 7	39%	0.50	39%	0.49	0.00	0.49	0.00	0.00
Has grade 8	26%	0.45	26%	0.44	0.00	0.44	0.00	0.00
Kindergarten cohort								
Cohort 2005	15%	0.36	15%	0.36	0.00	0.36	0.00	0.00
Cohort 2006	13%	0.37	13%	0.34	0.00	0.34	0.00	0.00
Cohort 2007	0%	0.00	0%	0.00	0.00	0.00		
Cohort 2008	13%	0.37	13%	0.34	0.00	0.34	0.00	0.00
Cohort 2009	14%	0.31	14%	0.35	0.00	0.34	0.00	0.00
Cohort 2010	13%	0.36	13%	0.33	0.00	0.34	0.00	0.00
Cohort 2011	16%	0.35	16%	0.37	0.00	0.37	0.00	0.00
Cohort 2012	15%	0.32	15%	0.36	0.00	0.36	0.00	0.00
Male	49%	0.50	49%	0.50	0.00	0.50	0.00	0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.04	0%	0.02	0.00	0.02	0.00	0.00
Asian	2%	0.10	2%	0.16	0.00	0.15	0.00	0.00
Black	1%	0.11	1%	0.10	0.00	0.11	0.00	0.00
Latinx	27%	0.41	27%	0.44	0.00	0.44	0.00	0.00

	Intervention (N = 1,161)		Control (N = 9,686)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Multiracial	6%	0.15	6%	0.24	0.00	0.24	0.00	0.00
Native Hawaiian/ Pacific Islander	0%	0.03	0%	0.02	0.00	0.02	0.00	0.00
White	63%	0.37	63%	0.48	0.00	0.47	0.00	0.00
Program participation								
English learner (ever)	19%	0.46	19%	0.39	0.00	0.40	0.00	0.00
Special education (ever)	2%	0.23	2%	0.13	0.00	0.15	0.00	0.00
Home language								
Spanish	19%	0.46	19%	0.39	0.00	0.40	0.00	0.00
Vietnamese	0%	0.06	0%	0.03	0.00	0.04	0.00	0.00
Other	1%	0.11	1%	0.10	0.00	0.10	0.00	0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A4. Baseline equivalency grade 3 reading and math outcomes

	Intervention (N = 842)		Control (N = 2,061)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	48.85	31.52	48.97	30.17	-0.12	30.56	0.00	
Kindergarten cohort								
Cohort 2010	15%	0.38	15%	0.36	0.00	0.36		0.00
Cohort 2011	0%	0.00	0%	0.00	0.00	0.00		
Cohort 2012	17%	0.38	17%	0.37	0.00	0.37		0.00
Cohort 2013	12%	0.31	12%	0.32	0.00	0.32		0.00
Cohort 2014	14%	0.36	14%	0.35	0.00	0.35		0.00
Cohort 2015	15%	0.34	15%	0.35	0.00	0.35		0.00
Cohort 2016	12%	0.29	12%	0.33	0.00	0.32		0.00
Home language								
English	0%	0.06	0%	0.05	0.00	0.05		0.00
Spanish	97%	0.15	97%	0.17	0.00	0.17		0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.06	0%	0.05	0.00	0.06		0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	3%	0.12	3%	0.16	0.00	0.15		0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.03	0%	0.02	0.00	0.03		0.00
Asian	0%	0.07	0%	0.06	0.00	0.06		0.00
Black	0%	0.03	0%	0.03	0.00	0.03		0.00
Latinx	97%	0.14	97%	0.17	0.00	0.16		0.00

	Intervention (N = 842)		Control (N = 2,061)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	2%	0.11	2%	0.15	0.00	0.14		0.00
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	53%	0.50	53%	0.50	0.00	0.50		0.00
Special education (ever)	6%	0.27	6%	0.24	0.00	0.25		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A5. Baseline equivalency grade 4 reading and math outcomes

	Intervention (N = 738)		Control (N = 1,705)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	47.03	31.07	47.25	30.29	-0.22	30.53	-0.01	
Kindergarten cohort								
Cohort 2009	17%	0.39	17%	0.38	0.00	0.38		0.00
Cohort 2011	17%	0.39	17%	0.38	0.00	0.38		0.00
Cohort 2012	19%	0.39	19%	0.39	0.00	0.39		0.00
Cohort 2013	14%	0.31	14%	0.35	0.00	0.34		0.00
Cohort 2014	16%	0.38	16%	0.37	0.00	0.37		0.00
Cohort 2015	16%	0.36	16%	0.37	0.00	0.37		0.00

	Intervention (N = 738)		Control (N = 1,705)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Home language								
English	0%	0.00	0%	0.00	0.00	0.00		
Spanish	97%	0.14	97%	0.17	0.00	0.16		0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.04	0%	0.04	0.00	0.04		0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	3%	0.13	3%	0.17	0.00	0.16		0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.04	0%	0.02	0.00	0.03		0.00
Asian	1%	0.06	1%	0.07	0.00	0.07		0.00
Black	0%	0.04	0%	0.03	0.00	0.04		0.00
Latinx	97%	0.14	97%	0.17	0.00	0.16		0.00
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	2%	0.11	2%	0.15	0.00	0.14		0.00
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	54%	0.50	54%	0.50	0.00	0.50		0.00
Special education (ever)	7%	0.28	7%	0.25	0.00	0.26		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A6. Baseline equivalency grade 5 reading and math outcomes

	Intervention (N = 607)		Control (N = 1,369)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	50.29	30.71	50.49	29.58	-0.21	29.93	-0.01	
Kindergarten cohort								
Cohort 2010	21%	0.41	21%	0.41	0.00	0.41		0.00
Cohort 2011	21%	0.42	21%	0.41	0.00	0.41		0.00
Cohort 2012	22%	0.41	22%	0.41	0.00	0.41		0.00
Cohort 2013	16%	0.33	16%	0.37	0.00	0.36		0.00
Cohort 2014	19%	0.41	19%	0.39	0.00	0.40		0.00
Home language								
English	0%	0.04	0%	0.03	0.00	0.03		0.00
Spanish	97%	0.14	97%	0.17	0.00	0.16		0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.04	0%	0.04	0.00	0.04		0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	3%	0.13	3%	0.17	0.00	0.16		0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.00	0%	0.00	0.00	0.00		
Asian	0%	0.06	0%	0.07	0.00	0.06		0.00
Black	0%	0.04	0%	0.03	0.00	0.03		0.00
Latinx	97%	0.13	97%	0.17	0.00	0.16		0.00
Multiracial	0%	0.00	0%	0.00	0.00	0.00		

	Intervention (N = 607)		Control (N = 1,369)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	2%	0.11	2%	0.16	0.00	0.14		0.00
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	53%	0.50	53%	0.50	0.00	0.50		0.00
Special education (ever)	7%	0.27	7%	0.25	0.00	0.25		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A7. Baseline equivalency grade 6 reading and math outcomes

	Intervention (N = 597)		Control (N = 1,325)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	47.03	31.07	47.25	30.29	-0.22	30.53	-0.01	
Kindergarten cohort								
Cohort 2009	21%	0.41	21%	0.41	0.00	0.41		0.00
Cohort 2010	20%	0.41	20%	0.40	0.00	0.41		0.00
Cohort 2011	21%	0.42	21%	0.41	0.00	0.41		0.00
Cohort 2012	21%	0.41	21%	0.41	0.00	0.41		0.00
Cohort 2013	17%	0.33	17%	0.37	0.00	0.36		0.00
Home language								
English	0%	0.06	0%	0.05	0.00	0.06		0.00

	Intervention (N = 597)		Control (N = 1,325)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Spanish	97%	0.14	97%	0.16	0.00	0.16		0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.04	0%	0.05	0.00	0.05		0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	2%	0.12	2%	0.15	0.00	0.14		0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.04	0%	0.03	0.00	0.03		0.00
Asian	1%	0.07	1%	0.08	0.00	0.08		0.00
Black	0%	0.00	0%	0.00	0.00	0.00		
Latinx	98%	0.13	98%	0.16	0.00	0.15		0.00
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	2%	0.11	2%	0.13	0.00	0.12		0.00
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	54%	0.50	54%	0.50	0.00	0.50		0.00
Special education (ever)	7%	0.26	7%	0.25	0.00	0.25		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A8. Baseline equivalency grade 7 reading and math outcomes

	Intervention (N = 509)		Control (N = 1,024)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	45.71	31.33	45.88	30.99	-0.17	31.10	-0.01	
Kindergarten cohort								
Cohort 2009	25%	0.43	25%	0.44	0.00	0.43		0.00
Cohort 2010	25%	0.43	25%	0.43	0.00	0.43		0.00
Cohort 2011	25%	0.44	25%	0.43	0.00	0.43		0.00
Cohort 2012	25%	0.43	25%	0.43	0.00	0.43		0.00
Home language								
English	0%	0.04	0%	0.03	0.00	0.04		0.00
Spanish	98%	0.12	98%	0.15	0.00	0.14		0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.04	0%	0.05	0.00	0.05		0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	2%	0.10	2%	0.14	0.00	0.13		0.00
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.04	0%	0.03	0.00	0.04		0.00
Asian	1%	0.06	1%	0.08	0.00	0.08		0.00
Black	0%	0.00	0%	0.00	0.00	0.00		
Latinx	98%	0.12	98%	0.15	0.00	0.14		0.00
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		

	Intervention (N = 509)		Control (N = 1,024)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
White	2%	0.09	2%	0.12	0.00	0.11		0.00
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	53%	0.50	53%	0.50	0.00	0.50		0.00
Special education (ever)	7%	0.25	7%	0.25	0.00	0.25		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A9. Baseline equivalency grade 8 reading and math outcomes

	Intervention (N = 378)		Control (N = 718)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	43.47	30.87	43.74	31.34	-0.27	31.18	-0.01	
Kindergarten cohort								
Cohort 2010	0.32	0.47	0.32	0.47	0.00	0.47		0.00
Cohort 2011	0.34	0.48	0.34	0.47	0.00	0.47		0.00
Home language								
English	0.00	0.05	0.00	0.05	0.00	0.05		0.00
Spanish	0.98	0.11	0.98	0.14	0.00	0.13		0.00
Chinese	0.00	0.00	0.00	0.00	0.00	0.00		
Vietnamese	0.00	0.05	0.00	0.06	0.00	0.06		0.00
Korean	0.00	0.00	0.00	0.00	0.00	0.00		
Other language	0.01	0.09	0.01	0.12	0.00	0.11		0.00

	Intervention (N = 378)		Control (N = 718)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Race/Ethnicity								
American Indian/ Alaska Native	0.00	0.00	0.00	0.00	0.00	0.00		
Asian	0.01	0.07	0.01	0.09	0.00	0.09		0.00
Black	0.00	0.00	0.00	0.00	0.00	0.00		
Latinx	0.98	0.10	0.98	0.13	0.00	0.12		0.00
Multiracial	0.00	0.00	0.00	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0.00	0.00	0.00	0.00	0.00	0.00		
White	0.01	0.07	0.01	0.10	0.00	0.09		0.00
Unknown race	0.00	0.00	0.00	0.00	0.00	0.00		
Male	0.53	0.50	0.53	0.50	0.00	0.50		0.00
Special education (ever)	0.07	0.26	0.07	0.25	0.00	0.26		0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A10. Baseline equivalency grade 11 reading and math outcomes

	Intervention (N = 158)		Control (N = 317)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Standardized ELPA RIT score in K	-0.34	0.68	-0.33	0.67	-0.01	0.67	-0.01	
Kindergarten cohort								
Cohort 2007	48%	0.50	48%	0.50	0.00	0.50	0.00	0.00
Cohort 2008	52%	0.50	52%	0.50	0.00	0.50	0.00	0.00
Home language								
English	0%	0.00	0%	0.00	0.00	0.00		
Spanish	100%	0.00	100%	0.00	0.00	0.00		
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	0%	0.00	0%	0.00	0.00	0.00		
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	0%	0.00	0%	0.00	0.00	0.00		
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.00	0%	0.00	0.00	0.00		
Asian	0%	0.00	0%	0.00	0.00	0.00		
Black	0%	0.00	0%	0.00	0.00	0.00		
Latinx	100%	0.00	100%	0.00	0.00	0.00		
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	0%	0.00	0%	0.00	0.00	0.00		

	Intervention (N = 158)		Control (N = 317)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Unknown race	0%	0.00	0%	0.00	0.00	0.00		
Male	51%	0.50	51%	0.50	0.00	0.50	0.00	0.00
Special education (ever)	3%	0.26	3%	0.18	0.00	0.20	0.00	0.00

RIT is Rasch Unit, dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A11. Baseline equivalency "on track" outcomes

	Intervention (N = 253)		Control (N = 510)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Woodcock RPI	40.18	30.91	40.61	31.29	-0.43	31.17	-0.01	
Kindergarten cohort								
Cohort 2008	0.02	0.15	0.02	0.15	0.00	0.15	0.00	0.00
Cohort 2009	0.47	0.50	0.47	0.50	0.00	0.50	0.00	0.00
Cohort 2010	0.51	0.50	0.51	0.50	0.00	0.50	0.00	0.00
Home language								
English	0.00	0.06	0.00	0.06	0.00	0.06	0.00	0.00
Spanish	0.99	0.11	0.99	0.12	0.00	0.11	0.00	0.00
Chinese	0.00	0.00	0.00	0.00	0.00	0.00		
Vietnamese	0.00	0.06	0.00	0.06	0.00	0.06	0.00	0.00
Korean	0.00	0.00	0.00	0.00	0.00	0.00		
Other language	0.01	0.06	0.01	0.08	0.00	0.07	0.00	0.00

	Intervention (N = 253)		Control (N = 510)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Race/Ethnicity								
American Indian/ Alaska Native	0.00	0.00	0.00	0.00	0.00	0.00		
Asian	0.00	0.06	0.00	0.06	0.00	0.06	0.00	0.00
Black	0.00	0.00	0.00	0.00	0.00	0.00		
Latinx	0.99	0.09	0.99	0.10	0.00	0.10	0.00	0.00
Multiracial	0.00	0.00	0.00	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0.00	0.00	0.00	0.00	0.00	0.00		
White	0.01	0.06	0.01	0.08	0.00	0.07	0.00	0.00
Unknown race	0.00	0.00	0.00	0.00	0.00	0.00		
Male	0.51	0.50	0.51	0.50	0.00	0.50	0.00	0.00
Special Education (ever)	0.07	0.23	0.07	0.26	0.00	0.25	0.00	0.00

dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table A12. Baseline equivalency high school graduation and college enrollment outcomes

	Intervention (N = 81)		Control (N = 196)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
English language proficiency								
Standardized ELPA RIT score in K	-0.31	0.67	-0.31	0.67	-0.01	0.67	-0.01	
Kindergarten cohort								
Cohort 2007	100%	0.00	100%	0.00	0.00	0.00		
Home language								
English	0%	0.00	0%	0.00	0.00	0.00		
Spanish	98%	0.16	98%	0.12	0.00	0.13	0.00	0.00
Chinese	0%	0.00	0%	0.00	0.00	0.00		
Vietnamese	2%	0.16	2%	0.12	0.00	0.13	0.00	0.00
Korean	0%	0.00	0%	0.00	0.00	0.00		
Other language	0%	0.00	0%	0.00	0.00	0.00		
Race/Ethnicity								
American Indian/ Alaska Native	0%	0.00	0%	0.00	0.00	0.00		
Asian	2%	0.16	2%	0.12	0.00	0.13	0.00	0.00
Black	0%	0.00	0%	0.00	0.00	0.00		
Latinx	98%	0.16	98%	0.12	0.00	0.13	0.00	0.00
Multiracial	0%	0.00	0%	0.00	0.00	0.00		
Native Hawaiian/ Pacific Islander	0%	0.00	0%	0.00	0.00	0.00		
White	0%	0.00	0%	0.00	0.00	0.00		
Unknown race	0%	0.00	0%	0.00	0.00	0.00		

	Intervention (N = 81)		Control (N = 196)		Baseline equivalency			
	mean	sd	mean	sd	diff	sdp	g	dcox
Male	52%	0.50	52%	0.50	0.00	0.50	0.00	0.00
Special education (ever)	4%	0.24	4%	0.19	0.00	0.20	0.00	0.00

RIT is Rasch Unit, dcox is Cox's index. diff is difference. g is Hedge's *g*. sd is standard deviation. sdp is pooled standard deviation.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Once students were matched, we followed students longitudinally and observed outcomes at different points of their K–12 and postsecondary trajectory using a regression analysis framework. We included the kindergarten variables in the analysis to control for any remaining variation between groups. English language arts and math outcomes from the Oregon state assessments were observed in grades 3–8, on track to graduation at the end of grade 9, high school graduation at the end of grade 12, and college enrollment at any point after the conclusion of grade 12. To estimate the impact of dual language participation on relevant outcomes, we used the following regression model:

$$Y_i = \beta_0 + \beta_1 DLI_i + X_i + \pi_c + \lambda_t + e_i$$

Where Y_i represents the assessment outcome for student (i) in the respective school year, DLI is a binary variable that indicates if a student participated in dual language for two or more years, X_i is the student baselined covariates (race/ethnicity, special education identification, gender, reported home language, initial English language proficiency screener score), π_c represents kindergarten entry cohort fixed effects, and λ_t represents year fixed effects. Standard errors were clustered at the school level to account for the nesting of students within schools. We included the Stata-produced analytic weights in the regression model. Our outcome of interest is represented by β_1, which represents the difference in assessment outcomes between English learner dual language students compared to English learner students not in a dual language program. To estimate the impact of dual language participation on the likelihood of being on track to graduate, graduate from high school, and enroll in any college, we used a similar approach but used logistic regression analysis.

Student growth trajectories by dual language participation

We used a multilevel model to estimate student growth trajectories and compare students who did and did not participate in dual language programs. We used CEM to match students on kindergarten demographic characteristics and initial English language proficiency scores in the same cohort entry year. In addition, we matched students on the grades that they were observed in the dataset. This means that a student who was observed in grades K–8 was matched with students who were also observed in those grades. Our growth model is represented by the following sets of equations:

Level 1

$$Y_{ti} = \pi_{0i} + \pi_1 Grade_{ti} + \pi_{2i} Grade_{2ti} + e_{ti}$$

Level 2

$$\pi_{0i} = \gamma_{00} + DLI\ Group_i + \mu_{0i}$$

$$\pi_{1i} = \gamma_{01} + DLI\ Group_i + \mu_{1i}$$

$$\pi_{2i} = \gamma_{02} + DLI\ Group_i$$

Where Y_{ti} represents the assessment outcome in time (t) for student (i). Grade represents the time variable, centered at 0 in kindergarten, and varies across students. Grade2 is the quadratic form of the Grade term. DLIGroup is a categorical variable that includes the following four categories: ever classified as English learner and in dual language, ever classified as English learner and not in dual language, never classified as English learner and in dual language, and never classified as English learner and never in dual language. To observe trajectories by group, DLIGroup is interacted with the Grade and Grade2 terms. Weights produced by the CEM matching approach were included in the model. All analyses were conducted using Stata 17 mixed command.

Time to reclassification

We used discrete-time survival analysis (Singer & Willet, 2003) to estimate likelihood of reclassification by dual language participation, similar to Thompson (2017). This approach answers how likely a specific event (reclassification from English learner services) will happen over a time period, takes into account individuals who never experience the event (were never reclassified), and uses student data up to the point when the event happens. As a first step in survival analysis, we estimate the hazard function which estimates the conditional probability that a student will be reclassified, assuming they have not already. This first step is performed using logistic regression and is represented by the following equation:

$$logit\ h(t_{it}) = \alpha_{1-13} + \beta_1 DLI_i + \beta_2 Grade_t + \beta_3 DLI_i x\ Grade_t + X_i + \pi_c + + e_{it}$$

Where *logit* $h(t_{it})$ is the conditional probability that the student will be reclassified in a specific grade and α_{1-13} represents a time period binary variable for each grade. DLI is a binary variable that indicates if a student participated in dual language for two or more years, $Grade_t$ is a continuous variable that represents time, $DLI_i \times Grade_t$ is the interaction between participating in dual language and the time variable, X_i are student-level covariates observed in kindergarten (race/ethnicity, special education identification, gender, reported home language, initial English language proficiency screener score), and π_c represents kindergarten entry cohort fixed effects.

The second step in the analysis is to estimate the survivor function. To do this, we used the estimates from the previous model to predict the cumulative proportion of students who were reclassified in each time period. The estimates can then be used to estimate the cumulative proportion of students who have not been reclassified after a particular number of years. The survivor function is represented by the following equation:

$$s(t_t) = s(t_{t-1})\,[1 - h(t_t)]$$

Where the survivor function for a time period (t) is estimated by multiplying the survival probability for a previous time period *(t – 1)* by one minus the hazard probability for the current time period. The hazard probability $[h(t_t)]$ is estimated from the previous model.

References

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Thompson, K. D. (2017). English learners' time to reclassification: An analysis. *Educational Policy, 31*(3), 330–363. (Online first, August 2015).

What Works Clearinghouse. (2015). *WWC standards brief.* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_baseline_080715.pdf

Appendix B. Data tables

Table B1. English language arts and math growth models

	Reading	Math
Ever English learner, not in dual language	-0.05	-0.09*
	(0.04)	(0.04)
Never English learner, in dual language	0.44***	0.33***
	(0.03)	(0.03)
Never English learner, not in dual language	0.46***	0.38***
	(0.03)	(0.03)
Grade	0.09***	0.09***
	(0.02)	(0.02)
Ever English learner, not in dual language X Grade	-0.02	-0.05*
	(0.02)	(0.03)
Never English learner, in dual language X Grade	-0.08***	-0.06**
	(0.02)	(0.02)
Never English learner, not in dual language X Grade	-0.09***	-0.09***
	(0.02)	(0.02)
Grade^2	-0.01***	-0.02***
	(0.00)	(0.00)
Ever English learner, not in dual language X Grade^2	-0.00	0.00
	(0.00)	(0.00)
Never English learner, in dual language X Grade^2	0.01	0.01**
	(0.00)	(0.00)
Never English learner, not in dual language X Grade^2	0.01*	0.01***
	(0.00)	(0.00)
Intercept	-0.33***	-0.27***
	(0.02)	(0.02)
Variance components		
Level 1 intercept	0.622***	0.607***
	(0.095)	(0.15)
Grade slope	0.021**	0.093**
	(0.006)	(0.07)

	Reading	Math
Level 1 error	-0.109***	0.096***
	(0.022)	(0.023)
Observations	**36,042**	**36,042**

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors in parentheses.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B2. English language arts models

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Dual language participation for 2+ years	0.04	0.08†	0.12†	0.17**	0.14*	0.15	0.20*
	(0.05)	(0.06)	(0.07)	(0.05)	(0.06)	(0.08)	(0.06)
Woodcock RPI English language proficiency screener	0.01***	0.01***	0.01***	0.01***	0.01***	0.00***	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
ELPA English language proficiency standard score							0.35*
							(0.13)
Spanish home language	-0.24		0.19	0.34*	-0.03	0.08	0.00
	(0.19)		(0.15)	(0.16)	(0.21)	(0.11)	(0.0)
Vietnamese home language	0.11	0.61	0.18	0.69*	0.30	0.52**	
	(0.17)	(0.36)	(0.15)	(0.25)	(0.24)	(0.16)	
Other home language	-0.15	0.05	0.15	0.50*	0.10	0.20	
	(0.19)	(0.11)	(0.18)	(0.21)	(0.22)	(0.15)	
Male	-0.07	-0.00	-0.01	-0.08	-0.06	-0.04	-0.08
	(0.03)	(0.05)	(0.08)	(0.07)	(0.09)	(0.09)	(0.05)
Special education status in kindergarten	-0.87***	-1.19**	-1.15*	-0.81**	-1.15*	-1.16*	-0.28**
	(0.22)	(0.37)	(0.42)	(0.26)	(0.47)	(0.43)	(0.08)
Constant	-0.23***	-0.19	-0.71**	-0.75**	0.14	0.11	-0.08
	(0.04)	(0.44)	(0.21)	(0.19)	(0.26)	(0.18)	(0.08)
Observations	**2,903**	**2,443**	**1,976**	**1,922**	**1,533**	**1,096**	**529**

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors in parentheses.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B3. English language arts improvement index

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Comparison group (Ever English learner, not in dual language)							
Sample size	2,061	1,705	1,369	1,325	1,024	718	371
Model-adjusted mean	-0.355	-0.329	-0.320	-0.310	-0.330	-0.321	-0.239
Standard deviation	1.078	1.126	1.173	0.968	1.037	0.975	0.865
Treatment group (Ever English learner, in dual language)							
Sample size	842	738	607	597	509	378	158
Model-adjusted mean	-0.320	-0.266	-0.196	-0.144	-0.190	-0.175	-0.085
Standard deviation	0.716	0.747	0.693	0.331	0.513	0.568	0.411
Difference	0.036	0.063	0.124	0.166	0.140	0.146	0.154
Pooled standard deviation	0.973	1.053	1.101	0.680	0.805	0.734	0.576
Omega	1.000	1.000	1.000	1.000	1.000	0.999	0.999
Effect size	0.037	0.059	0.112	0.245	0.174	0.198	0.266
Model estimated *p*-value	0.505	0.164	0.083	0.004	0.028	0.103	0.013
Improvement index	**1.5**	**2.4**	**4.5**	**9.7**	**6.9**	**7.9**	**10.5**

Note: Effect sizes for continuous variables show results from Hedge's *g*. Effect sizes for binary variables show results from Cox's index.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B4. Math models

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Dual language participation for 2+ years	0.06† (0.06)	0.15* (0.06)	0.17* (0.08)	0.17* (0.06)	0.20* (0.08)	0.13† (0.08)	0.20** (0.06)
Woodcock RPI English language proficiency screener	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.00*** (0.00)	
ELPA English language proficiency standard score							0.36* (0.12)
Spanish home language	-0.00 (0.11)		0.26 (0.15)	0.19 (0.13)	0.06 (0.16)	0.17 (0.12)	0.00 (0.0)

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Vietnamese home language	0.37***	0.38*	0.52**	0.60**	0.56*	0.76***	
	(0.08)	(0.17)	(0.17)	(0.20)	(0.22)	(0.16)	
Other home language	0.10	0.08	0.32	0.44*	0.26	0.65**	
	(0.10)	(0.09)	(0.18)	(0.17)	(0.17)	(0.20)	
Male	0.04	0.10	0.08	-0.01	0.03	0.02	-0.03
	(0.03)	(0.06)	(0.08)	(0.06)	(0.09)	(0.08)	(0.05)
Special education status in kindergarten	-0.82**	-1.16**	-1.10*	-0.80**	-1.12*	-1.15*	-0.24*
	(0.24)	(0.39)	(0.43)	(0.25)	(0.44)	(0.39)	(0.08)
Constant	-0.46***	-0.26*	-0.78***	-0.68***	-0.00	0.22	-0.12
	(0.03)	(0.12)	(0.21)	(0.15)	(0.23)	(0.20)	(0.06)
Observations	**2,903**	**2,443**	**1,976**	**1,922**	**1,533**	**1,096**	**529**

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors in parentheses.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B5. Math improvement index

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Comparison group (Ever English learner, not in dual language)							
Sample size	2,061	1,705	1,369	1,325	1,024	718	371
Model-adjusted mean	-0.337	-0.349	-0.329	-0.363	-0.402	-0.366	-0.257
Standard deviation	1.022	1.122	1.143	0.955	1.019	0.937	0.823
Treatment group (Ever English learner, in dual language)							
Sample size	842	738	607	597	509	378	258
Model-adjusted mean	-0.274	-0.196	-0.163	-0.192	-0.204	-0.231	-0.097
Standard deviation	0.777	0.704	0.704	0.385	0.540	0.623	0.421
Difference	0.063	0.153	0.166	0.170	0.197	0.135	0.160
Pooled standard deviation	0.916	1.029	1.057	0.675	0.790	0.709	0.473
Omega	1.000	1.000	1.000	1.000	1.000	0.999	0.999
Effect size	0.069	0.148	0.157	0.252	0.250	0.190	0.338

	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
Model estimated *p*-value	**0.296**	**0.016**	**0.049**	**0.015**	**0.021**	**0.128**	**0.009**
Improvement index	**2.7**	**5.9**	**6.3**	**10.0**	**9.9**	**7.5**	**13.2**

Note: Effect sizes for continuous variables show results from Hedge's *g*. Effect sizes for binary variables show results from Cox's index.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B6. Hazard analysis model

	Percentage of students predicted to be reclassified at the beginning of each grade			Cumulative percentage of students predicted to be reclassified at the beginning of each grade		
Grade	**All students**	**Comparison**	**Dual language**	**All students**	**Comparison**	**Dual language**
K	0%	0%	0%	0%	0%	0%
1	7%	7%	4%	7%	7%	4%
2	11%	12%	8%	16%	17%	11%
3	15%	15%	12%	27%	28%	20%
4	23%	23%	20%	41%	42%	34%
5	35%	35%	35%	57%	58%	53%
6	44%	44%	47%	71%	72%	70%
7	44%	43%	49%	80%	80%	80%
8	34%	32%	41%	83%	83%	85%
9	28%	26%	38%	86%	85%	88%
10	19%	17%	30%	87%	86%	90%
11	26%	23%	40%	88%	87%	92%
12	3%	2%	7%	88%	88%	92%
Observations	**7,783**	**6,453**	**1,330**	**7,783**	**6,453**	**1,330**

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B7. Graduation and postsecondary models

	On track to graduation (grade 9)	High school graduate (ever)	Enrolled in college (ever)
Dual language participation for 2+ years	1.37*	1.23	1.34
	(0.19)	(0.54)	(0.42)
Woodcock RPI English language proficiency screener	1.00		
	(0.00)		
ELPA English language proficiency standard score		1.58*	1.94***
		(0.32)	(0.29)
Spanish home language	0.35		
	(0.38)		
Male	0.62***	0.41***	0.61**
	(0.08)	(0.09)	(0.11)
Special education status in kindergarten	1.41	4.83	4.49*
	(0.34)	(6.02)	(2.99)
Cohort 2009	6.5e+05***		
	(6.3e+05)		
Cohort 2010	1.1e+06***		
	(1.1e+06)		
Observations	**763**	**277**	**277**

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Standard errors in parentheses; exponentiated coefficients.

Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Table B8. Graduation and postsecondary improvement index

	On track to graduation	Graduation	Postsecondary enrollment
Comparison group (Ever English learner, not in dual language)			
Sample size	510	196	196
Model-adjusted mean	0.715	0.811	0.526
Standard deviation	0.445	0.388	0.501
Treatment group (Ever English learner, in dual language)			
Sample size	253	81	81
Model-adjusted mean	0.773	0.840	0.593

	On track to graduation	Graduation	Postsecondary enrollment
Standard deviation	0.421	0.369	0.494
Difference		0.029	0.067
Pooled standard deviation	0.191	0.147	0.249
Omega	0.999	0.997	0.997
Effect size	0.185	0.120	0.164
Model estimated *p*-value	0.027	0.633	0.346
Improvement index	**7.3**	**4.8**	**6.5**

Note: Effect sizes for continuous variables show results from Hedge's *g*. Effect sizes for binary variables show results from Cox's index.

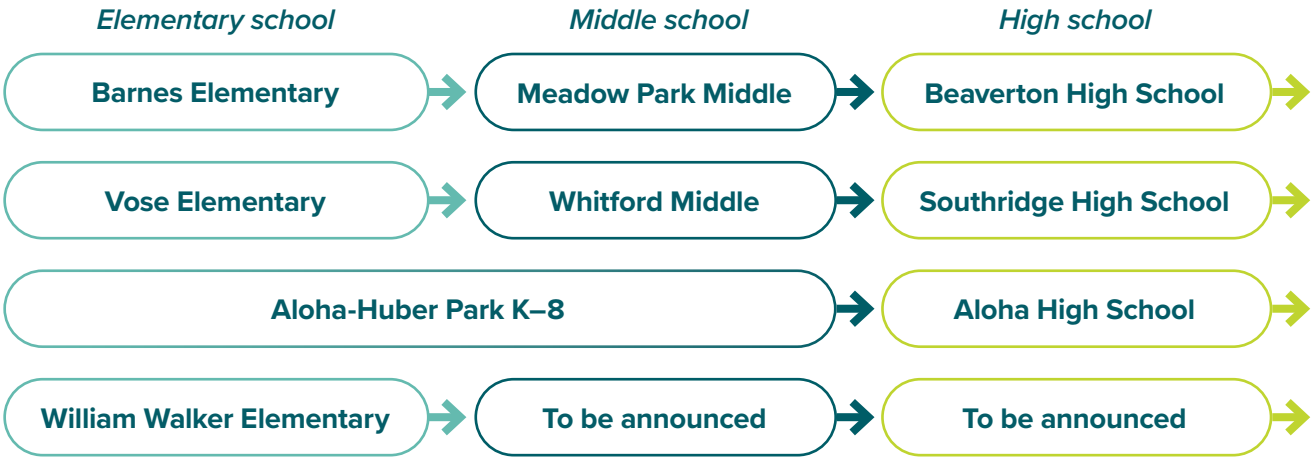
Source: Education Northwest analysis of Beaverton School District administrative data, 2006–2019.

Appendix C. History of Beaverton School District dual language program and program models

> *"The mission of the Beaverton School District dual language program is to honor and develop multilingual, multi-literate, and multicultural students through rigorous, culturally inclusive education while nurturing a diversity of identities, and empowering students to become agents of change in a global community." (link)*

The Beaverton School District dual language program began in the early 1990s as a grassroots effort of parents, community members, and educators at Barnes Elementary School. Thirty years later, in fall 2023, Spanish-English dual language classes were offered at three elementary schools, two middle schools, one K–8 school, and three high schools in the district (figure C1). This count includes William Walker Elementary School that began two dual language kindergarten strands in fall 2023. Data from William Walker are not included in this study.

Figure C1. Dual language school pathways in Beaverton School District, fall 2023



Source: Beaverton School District Dual Language Website (2024).

Beaverton School District has announced that three new dual language programs will begin in fall 2024. These include a Spanish-English dual language pre-kindergarten strand at McKinley Elementary School, a Spanish-English kindergarten strand at Hazeldale Elementary School, and a Mandarin Chinese-English kindergarten strand at Jacob Wismer Elementary School.

Beaverton dual language schools use 90:10, 80:20, and 50:50 models. In 90:10 models, 90 percent of the instructional day in kindergarten is in the target language, with increasing amounts of English instruction each grade. Some research has found that 90:10 models result in greater oral and literacy attainment in Spanish among both English learners and home language English speakers (Lindholm-Leary, 2020). On the other hand, 50:50 models provide equal language allocation regardless of grade level (table C1).

Table C1. Beaverton School District dual language school models and histories

School	History
Barnes Elementary School (1992)	Barnes began as a 50:50 program model in fall 1992 with native language literacy instruction, meaning that students are initially taught to read in their home language. Barnes switched to a 90:10 model in 2007, with all participating students learning to read in Spanish first. This change was made in response to student academic and English language performance in the upper grades and is backed by research. Barnes began transitioning from a strand to schoolwide program in 2019 and is expected to be a fully implemented pre-K–5 dual language school by fall 2024 as part of an expanded zone with two neighboring elementary schools. The program began transitioning to an 80:20 model in the 2023–24 school year. English language development is offered as an integrated co-teaching service model with limited and targeted pull-out.
Vose Elementary School (2001)	Vose began as a 50:50 program model in fall 2001 with native language literacy instruction. It is a strand school with three dual language classrooms at each grade. In 2004, in response to research recommendations, Vose switched to a 90:10 model with all participating students learning to read in Spanish first. English language development is offered as an integrated co-teaching service model with limited and targeted pull-out. The program began transitioning to an 80:20 model in the 2023–24 school year.
Aloha-Huber Park K-8 (2000)	Aloha-Huber Park began as an elementary 50:50 maintenance model in fall 2000 to strengthen native language for Spanish-speaking populations, and to provide an opportunity to learn Spanish for home language English speakers. In 2021 the school began transitioning to a 90:10 model. Aloha-Huber Park is a strand school, with two dual language classrooms at each K–5 grade. English language development is offered as an integrated co-teaching service model with limited and targeted pull-out in K–5. In grades 6-8 dual language instruction began in fall 2006, as a self-contained 50:50 model. The program developed from community request to continue supporting academic achievement for students participating in the elementary program. In 2011 the delivery of bilingual instruction shifted to language arts fully in Spanish. Currently, English language development is provided through an integrated co-teaching model in grades 6–8. In fall 2024, Aloha-Huber Park will expand to three strands beginning at the kindergarten level and will transition to a schoolwide dual language program at the middle school level.

School	History
Meadow Park Middle School (1998)	Meadow Park Middle School has offered dual language instruction on and off since 1998 as a continuation of Barnes Elementary School's dual language program. As of 2023 all dual language students at Meadow Park take integrated Spanish language arts, science, and math in Spanish. English language development is delivered through an integrated co-teaching model and an ELD class period is offered for newcomer students and students testing at the emerging level.
Whitford Middle School (2006)	Whitford Middle School began its dual language program in fall 2006 with a grade 5 teacher from Vose Elementary looping up to provide a Spanish-language humanities class. In 2024 Whitford Middle School offers bilingual language arts and dual language science to all dual language program students in grades 6–8. Math in Spanish has also been offered. English language development is delivered through an integrated model and an ELD class period is offered for newcomer students and students testing at the emerging level.
Beaverton High School (2015)	Beaverton High School began its dual language program in 2015, continuing the pathway from Meadow Park Middle School and Barnes Elementary School. Each year, students participating in the dual language program take a Spanish-language arts course, culminating in an Advanced Placement Spanish language and literature course. They also take at least one other core course in Spanish—for example, world and U.S. history, physics and chemistry, and up to three years of integrated algebra, geometry, and statistics. English language development is delivered during a class period for all students classified as English learners.
Southridge High School (2001)	Southridge High School began its dual language program in 2001, continuing the pathway from Whitford Middle School and Vose Elementary School. Each year, dual language students take Spanish language arts, including an International Baccalaureate Spanish language and literature course. They also take at least one other core class in Spanish: U.S. history and civics/economics, physics, chemistry, International Baccalaureate biology, and up to three years of integrated algebra, geometry, and statistics. English language development is delivered during a class period for most English learner classified students awhile some students participate in an integrated ELD and content model.
Aloha High School (2009)	Aloha High School began its dual language program in 2009, continuing the pathway from Aloha-Huber Park K–8 School. Each year, dual language students take Spanish language arts, including both dual credit and Advanced Placement courses. They also take at least one other core class in Spanish. These include U.S. history and global studies as well as physics and chemistry. English language development is delivered through an integrated ELD and content model and an ELD class period is offered for newcomer students and students testing at the emerging level.

References

Beaverton School District. (2016). *ELL program road maps: Dual language.* Prepared by Education Northwest for Beaverton School District. https://www.cosa.k12.or.us/sites/default/files/images/ell_program_roadmaps_intro.pdf

Lindholm-Leary, K. (2020) *Current research findings on two-way bilingual immersion Education.* San Jose State University. https://atdle.org/wp-content/uploads/2020/07/lindholmleary2020_current-research-twbi-3.pdf