



NORTH SHORE SCHOOL DISTRICT 112

Dual Language Program Evaluation Final Report

An external evaluation of the K–5 two-way Spanish–English dual language program, prepared as the district plans for program consolidation in 2027–28.

CONTRACTED DELIVERABLES

- 01 Class Size Recommendations
- 02 Curriculum Alignment, Coherence & Pacing Recommendations
- 03 Biliteracy Assessment Framework Recommendations
- 04 Program Model Recommendations

PREPARED BY

John Hilliard, MEd
Project Director

Margo Gottlieb, PhD
Senior Evaluation Consultant

Cristina Sánchez-López, MEd
Co-Principal Investigator and On-Site Coordinator

Fred Genesee, PhD
Dual Language Researcher and Data Analyst



Self-portraits by students in the SD 112 dual language program.

Table of Contents

List of Figures and Tables	iii
Executive Summary	v
I. Introduction & Context	1
II. Research Foundations	7
III. Design for Dual Language Evaluation	12
IV. Overall Methodology & Data Sources	16
V. Data Analyses, Reporting & Findings	24
V.1 · EQ 1: English Language Proficiency (ELP)	25
V.2 · EQ 2: Spanish Language Proficiency (SLP)	28
V.3 · EQ 3: Academic Achievement	31
V.4 · EQ 4: Biliteracy Assessment Framework	37
V.5 · EQ 5: Class Size and Composition	40
V.6 · EQ 6: Curriculum Coherence and Pacing	47
VI. Deliverables and Recommendations	52
VI.1 · Class Size Recommendations	52
VI.2 · Curriculum Alignment, Coherence & Pacing Recommendations	57
VI.3 · Biliteracy Assessment Framework Recommendations	64
VI.4 · Program Model Recommendations	73
VII. Prioritized Recommendations for Program Implementation	80
VIII. Limitations of the Evaluation Study	84
Glossary	86
References	93
Appendices	95
Appendix A Evaluation Question Maps/ Guiding Principles Crosswalk	95
Appendix B Data Collection Instruments	102
Appendix C Interview and Focus Group Protocols	105
Appendix D Proposed Year-Long Biliteracy Assessment Schedule	109
About the Authors	111

LIST OF FIGURES AND TABLES

I. INTRODUCTION & CONTEXT

- Figure I.1** North Shore School District 112 enrollment and demographic trends, 2020–21 through 2025–26
- Figure I.2** Report structure at a glance

III. DESIGN FOR DUAL LANGUAGE EVALUATION

- Figure III.1** The Dual Language Program Evaluation Blueprint: the seven EQs leading to the four RFP deliverables
- Figure III.2** Anatomy of a Map of an Evaluation Question

IV. OVERALL METHODOLOGY & DATA SOURCES

- Table IV.1** Evidence base at a glance

V.1 EQ 1: ENGLISH LANGUAGE PROFICIENCY

- Figure V.1.1** English language proficiency: what the ACCESS data show

V.2 EQ 2: SPANISH LANGUAGE PROFICIENCY

- Figure V.2.1** DALE proficiency-level estimate and district SLA rubric, by writing sample

V.3 EQ 3: ACADEMIC ACHIEVEMENT

- Table V.3.1** Growth and status on NWEA MAP in English language arts and mathematics in English, three years
- Table V.3.2** The same measures by grade, 2024–25
- Figure V.3.3** One cohort followed for three years
- Figure V.3.4** IAR Spring 2025, dual language program
- Figure V.3.5** What teachers report, by theme
- Figure V.3.6** The review of curricular materials
- Figure V.3.7** The English Learner strand: what the evidence does and does not support

V.4 EQ 4: BILITERACY ASSESSMENT FRAMEWORK

- Figure V.4.1** The measures in use, K–5, by area, language, and purpose

V.5 EQ 5: CLASS SIZE AND COMPOSITION

- Table V.5.1** The program enumerated, 2025 distribution
- Figure V.5.2** Average class size and EL/LEP share by grade and site, 2025 distribution
- Figure V.5.3** EL and not-EL students by grade and site, 2025 distribution

V.6 EQ 6: CURRICULUM COHERENCE AND PACING

- Table V.6.1** Core Literacy Module start month by grade, SY 2025–26
- Table V.6.2** Literacy materials evaluation: overall and dual language-specific component ratings
- Table V.6.3** Mathematics materials evaluation: overall and dual language-specific component ratings

VI.1 CLASS SIZE RECOMMENDATIONS

- Table VI.1.1** Sections required at the recommended class size of 20 and at the upper limit of 24

VI.2 CURRICULUM & PACING FRAMEWORK RECOMMENDATIONS

- Figure VI.2.1** A Proposed Dual Language Biliteracy Curricular Framework
- Figure VI.2.2** Suggested implementation map for the recommendations

VI.3 BILITERACY ASSESSMENT FRAMEWORK RECOMMENDATIONS

- Figure VI.3.1** A Proposed Dual Language Biliteracy Curricular Framework onto which the Biliteracy Assessment Framework is superimposed
- Figure VI.3.2** A Proposed Dual Language Biliteracy Assessment Framework
- Table VI.3.3** Suggested pool of assessment instruments by purpose and language
- Table VI.3.4** Areas of assessment currently in place and suggested instruments
- Figure VI.3.5** Considerations for annual and interim instruments and measures for Spanish language proficiency, Spanish language arts, and mathematics

VI.4 PROGRAM MODEL RECOMMENDATIONS

Figure VI.4.1 The program-model recommendation: evidence from V.1–V.6 against the three pillars

APPENDIX A: EVALUATION QUESTION MAPS

EQ 1 Map	English Language Proficiency (ELP)
EQ 2 Map	Spanish Language Proficiency (SLP)
EQ 3 Map	Academic Achievement
EQ 4 Map	Biliteracy Assessment Framework
EQ 5 Map	Class Size
EQ 6 Map	Curriculum and Pacing Framework
EQ 7 Map	Program Model

APPENDIX B: DATA COLLECTION INSTRUMENTS

- Qualitative Data Tracker
- Quantitative Data Tracker
- Instructional Materials Review Protocol
- Dual Language Classroom Observation Rubric
- Assessment Inventory: Verification Worksheet
- 50/50 Model Schedule, Kindergarten (Weeks 1–2)

APPENDIX C: INTERVIEW AND FOCUS GROUP PROTOCOLS

- Administrative Staff Interview Protocol (front office staff)
- Engagement Specialist Interview Protocol
- Parent Interview Field Protocol (English)
- Entrevista a Padres: Protocolo (Spanish)
- Stakeholder Interview Protocols (cover page)
- Student Focus Group Field Protocol
- Teacher Focus Group Field Protocol

EXECUTIVE SUMMARY

North Shore School District 112 commissioned an evaluation of its K–5 two-way dual language program at Oak Terrace and Red Oak Schools. Paridad Education Consulting conducted the evaluation study from March through August 2026, answering seven evaluation questions based on the Three Pillars of Dual Language Education: Bilingualism and Biliteracy, High Academic Achievement, and Socio-cultural Competence. This study aims to provide a comprehensive report that puts forth findings and recommendations per the district's four deliverables as identified in the December 2025 Request for Proposals.

Research Foundation

This evaluation report sets out the research and theory behind the evaluation of School District 112's K–5 dual language (DL) program. It synthesizes peer-reviewed research, large-scale longitudinal studies, and professional guidance documents on two-way dual language education, with particular attention to studies of 50% English/50% Spanish program models comparable to SD 112's design. The review is intended to support professional judgment about SD 112's specific context.

Evaluation Design

The evaluation is based on the Guiding Principles for Dual Language Education and in the three pillars of dual language education: high academic achievement, bilingualism and biliteracy, and sociocultural competence. The three pillars are named in the district's RFP as the framework against which program effectiveness should be assessed.

The findings are organized around six primary evaluation questions (EQs) and a seventh overarching question: EQ 1 (English language proficiency), EQ 2 (Spanish language proficiency), EQ 3 (academic achievement), and EQ 4 (biliteracy assessment), EQ 5 (class size and composition), EQ 6 (curriculum and pacing) and a seventh overarching question, EQ 7, which addresses the program model and is analyzed within the instructional model deliverable.

The evidence base for the evaluation draws from quantitative and qualitative district-specific data, research from the field of dual language education, and the cumulative knowledge and professional experience of the Paridad Team. Some of the data analyzed in this report include: district assessment data, classroom observations, administrator interviews, teacher, parent and student focus groups, a district assessment inventory, the district's pacing schedule, and curriculum documents.

Building from Strengths

The district's long established dual language program has strong foundational elements in place. During the evaluation process, the team noted strengths and resources reflected in both the qualitative and quantitative evidence. The language proficiency and academic achievement data reveal areas of significant growth from which to make any of the suggested recommendations contained in this report. The classroom observations, focus group conversations, and materials review portray a district community with a commitment to maintain and improve their dual language program and with the materials and resources to adequately support it. Evaluations tend to focus on areas in need of improvement. This report also marks where SD 112's dual language program is working: where students are doing well, or better than a stated comparison, and where the district already has strengths the recommendations can build on. We call these bright spots (Heath & Heath, 2010). Each one is marked wherever it appears, so a reader can follow it from the summary to the evidence. The ✨ means the evidence supports a strength, although the area may still need attention. Several bright spots are listed beside findings that may point to an area that does need attention, and the report names them together.

Dedicated Teaching Staff

In focus group sessions and individual interviews, school and district personnel clearly expressed their commitment to the success of students in the dual language program. Classroom observations reveal teachers who implement the curriculum given to them. Their interactions with students were consistently positive, supportive and respectful with an emphasis on establishing safe learning environments for their classrooms and throughout the schools. Observers also noted instructors with high levels of proficiency in both English and Spanish.

Family & Student Endorsement of the Program

The family members, who participated in the focus groups, express ardent belief in dual language education; their dedication to and support of the program for all students were clearly communicated. For them, the benefits of language and cultural learning are very clear, and maintenance of the dual language education program is a priority. This sentiment is echoed in the student focus groups, as they embrace the opportunity to become bilingual and biliterate as captured in the conversations, in both languages, that emerged during the student focus groups.

English Language Proficiency for English Learners

In a program committed to building bilingualism and biliteracy for all learners, it is clear that students' development of two languages does not hinder their growth in the majority language, English. In fact, English learners in the dual language schools made sustained gains in English language proficiency. The 119 students tested in all three springs moved from a mean ACCESS composite scale score of 258 to 335. Scale score growth was also evident across modalities, both for oral language (listening and speaking) and literacy (reading and writing). By 2024–25, 19.3 percent of the cohort had reached the Illinois reclassification criterion of 4.8. Benchmarked against WIDA Consortium averages, three of the four grade cohorts met or exceeded the matching consortium gain across the two-year window.

Academic Achievement

Academic achievement in English follows the same pattern of growth of English proficiency. On the NWEA MAP test, dual language students grow faster than national peers who began at the same score, in all three years in both mathematics and language arts in English. In the Spring 2025 IAR administration, the dual language program students exceeded the state average in every grade and subject area .

Areas of Emerging Growth and Promise: English and Spanish Development

Another area of promise is the collaborative professional learning offered by *With Partners*. Throughout the teacher focus groups educators express appreciation for the dual language-specific professional learning ***With Partners*** conducted in both languages. Their work on aligning and implementing the bilingual writing curriculum has been received positively.

Additionally, the Paridad team notes that the dual language program has a rich compendium of resources in both languages. During observed instruction, teachers demonstrate classroom management, cooperative learning, modeling, and student engagement strategies. Several staff members were identified as potential teacher leaders and content specialists who could share their practices with one another to strengthen the multilingual and multicultural practices of the entire teaching staff.

Summary of Recommendations According to RFP Deliverables

This report is tasked with providing recommendations under four categories: optimal class sizes, curriculum pacing and coherence, a biliteracy assessment framework, and program language

allocation model. The following recommendations should be taken in the light of an upcoming (December 2026) release of an Illinois State Board of Education (ISBE) guidance document for dual language programs that may provide further recommendations with input from stakeholders from across the state of Illinois.

Class Size Recommendations

A review of the research base around this question points to instructional quality as a larger influence on student performance than actual class size. In light of this research and Illinois State Board of Education School Code, this report recommends a target of 20 students per section, with a ceiling of 24 and at least two sections at every grade. Due to natural EL status shifts as students progress in the program, we recommend that the composition of the classrooms should be set with regard to the potential for both linguistically homogenous and heterogenous groupings at each grade, rather than as a set program-wide ratio. The current program is functioning well below these recommended size thresholds and already meets the Illinois regulatory ceiling set in (23 Ill. Adm. Code 228.30(b)(1)). In both cases, the dual language program class sizes are well within those parameters with most classes having an average of 18 students. The forthcoming ISBE guidance, noted above, should shed further light on regulatory targets for class size in dual language programs.

Curriculum and Pacing Recommendations

The existing literacy pacing guide, created to coordinate instruction in two languages, provides a clear roadmap for biliteracy development, and balances English and Spanish instructional time. However, the team found that while the English and Spanish content area curriculum materials (language arts, science, math and social studies) are abundant and overall of high quality, implementing the many components of each of the two monolingual curricula require more time than is allotted in the current schedule.

Accordingly, the Paridad team recommends the following: 1) convene collaborative practitioner/administrator teams to work on reviewing each subject area's curriculum with the aim of knitting the two language curricula together rather than trying to teach all content twice in each language. The collaborative teams would weave in cross-linguistic and cross-cultural connections throughout each topic or unit of study (in each content area), so that the new maps would represent a dual language model that is cohesive between the two languages rather than two monolingual curricula; 2) maintain the current alternating week language arts schedule by full week or full module with opportunities for planned, evaluable cross-linguistic transfer throughout the entire week.

Biliteracy Assessment Framework Recommendations

The conceptualization of a biliteracy assessment framework is one that demonstrates a robust strong presence of Spanish and English in language proficiency and language arts. Parallel measures in both languages and the interaction between the two would allow for the creation of a coordinated and comprehensive system. The display of longitudinal data reveal student performance across languages over time. To strengthen the framework, more Spanish data are needed to provide a balanced and stable portrait of student growth over time. The list of recommendations below for conceptualizing a biliteracy assessment framework is extensive, however its intent is to offer a full spectrum of possibilities from which the district can select, pilot, and adopt as it feels would be most effective in meeting its dual language education goals.

Recommendations include: 1) Build a balanced assessment framework organized by purpose and language; 2) invite teams of dual language teachers to design and pilot common biliteracy measures based on curricular units of instruction (e.g., from language arts, science, social studies) across classrooms to complement standardized testing data; 3) create and archive a biliteracy portrait for every dual language student that includes test data, student goals, and self-reflection; 4) place the selection of assessment measures created by a team of teachers on a

year-long calendar, and evaluate the effectiveness of the biliteracy assessment framework annually.

Dual Language Program Model Recommendation

Qualitative and quantitative data sources converge on the recommendation for an optimal dual language education program model that reconfigures delivery of the current program to achieve a true 50/50 language allocation over the next two-three academic years. This recommendation is an intermediary step in answering the overarching question of this report which is whether to maintain a 50/50 model or adopt a 80/20 model.

Limitations of this Evaluation

This is a descriptive evaluation that did not use control or experimental groups. Therefore, every recommendation, informed by multiple sources of evidence, also includes professional judgment from the team in addition to research from the field of bilingual and dual language education. Classroom observations covered two weeks in April 2026, making it difficult to generalize findings across the school year. Additionally, Spanish language proficiency and achievement analyses are based on a small set of writing samples and on one interim instrument, MAP Reading in Spanish. Sociocultural competence, the third pillar of the dual language guiding principles was not explicitly evaluated in order to not interfere with the district's ongoing initiatives and established processes in this area.

Closing Statement

The diligence and cooperation shown by the district leadership, support staff, teachers, parents and students have been outstanding throughout the entire evaluation process. Their insights and the historical knowledge have contributed to the strength and validity of this report.

I. INTRODUCTION & CONTEXT

Purpose and Scope of the Evaluation

North Shore School District 112 issued a Request for Proposals in December 2025 seeking a qualified external consultant to conduct a comprehensive evaluation of its 50/50 Dual Language Program. Paridad Education Consulting submitted a proposal on January 23, 2026, revised February 10, 2026. Following a competitive process the Board of Education approved the selection of the consultant on February 18, 2026, and the evaluation commenced on March 2, 2026.

The district framed the purpose of the evaluation as assessing program effectiveness, identifying strengths and areas for growth, and producing actionable recommendations to strengthen student outcomes in both Spanish and English across participating grade levels. The RFP directed that the evaluation assess the program's effectiveness in promoting the three pillars of dual language education (high academic achievement, bilingualism and biliteracy, and sociocultural competence) and evaluate the current instructional model.

The evaluation is contracted to produce four deliverable sets of recommendations in addition to the findings themselves: an instructional model analysis and recommendations, a set of recommendations for a comprehensive biliteracy assessment framework, a set of recommendations for a curriculum and pacing framework, and a class size analysis and recommendations. A crosswalk locating each deliverable within this report appears in the Organization of the Report, located at the end of this section.

What this evaluation is. A comprehensive mixed-methods program evaluation, based on the Guiding Principles for Dual Language Education (3rd edition, Center for Applied Linguistics, 2018) and referencing the Prism Model (Collier, Thomas & Lachance, 2026). It combines analysis of multi-year student outcome data in both languages with structured classroom observation, stakeholder interviews with teachers, administrators, families, and students, survey data, and document review.

Limits of the design. The evaluation is observational. Students were not assigned to the dual language program at random, and families who choose the program may differ systematically from those who did not. Comparisons between program participants and non-participants are descriptive and should not be read as estimates of the program's causal effect.

The evaluation period ran from March 2026 through August 2026.

District and Program Context

North Shore School District 112 serves the communities of Highland Park, Highwood, and Fort Sheridan, Illinois. The district enrolled 3,695 students across nine schools in 2024–25, the most recent year reported on the Illinois Report Card. The district's December 2025 Request for Proposals cited an enrollment of 3,752.

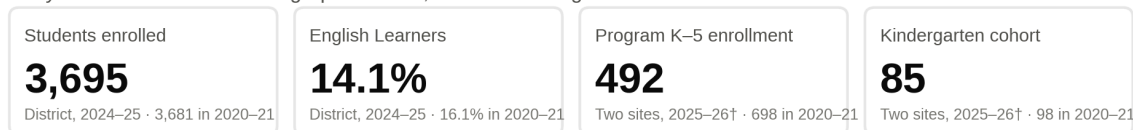
District demographics for 2024–25, the most recent year published on the Illinois Report Card, are as follows. English Learners constituted 14.1 percent of district enrollment, below the state figure of 17.5 percent. Students with Individualized Education Programs constituted 16.6 percent, essentially at the state rate of 16.3 percent. Low-income students constituted 20.4 percent, well below the state figure of 49.7 percent. By racial and ethnic composition, the district enrolled a substantially larger share of White students than the state (67.8 percent against 44.3 percent) and smaller shares of Hispanic or Latino students (21.9 percent against 28.6 percent) and Black students (1.8 percent against 16.3 percent).

Figure I.1 places these indicators in a six-year series, from 2020–21 through 2025–26. Over that period, district totals changed little, while program-level figures changed considerably. District enrollment moved from 3,681 to 3,695 between 2020–21 and 2024–25, a change of less than half a percent, and the district English Learner share declined from 16.1 percent to 14.1 percent. Over the same years, K–5 enrollment at the two dual language sites fell from 698 to 523, and stood at 492 in 2025–26, a decline of 29.5 percent across the six years, concentrated at Oak Terrace, where K–5 enrollment fell from 440 to 293. The entering kindergarten cohort at the two sites fell from 98 to 80 across the five published years and stood at 85 in

2025–26, distributed as three sections at Oak Terrace and one at Red Oak. Site English Learner shares moved less than the district figure: Oak Terrace from 45.2 to 44.0 percent, Red Oak from 36.8 to 29.1 percent. Figures for 2025–26 are drawn from the district classroom roster collected during the April 2026 observation visits; the Illinois Report Card publishes 2025–26 in late October 2026.

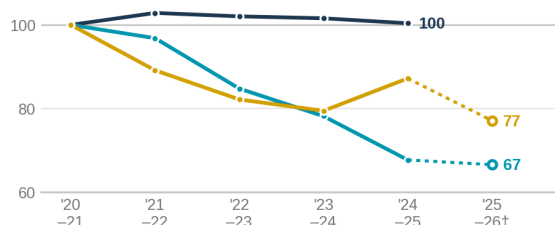
North Shore School District 112 · Student Demographics

Six-year enrollment and demographic trends, 2020–21 through 2025–26

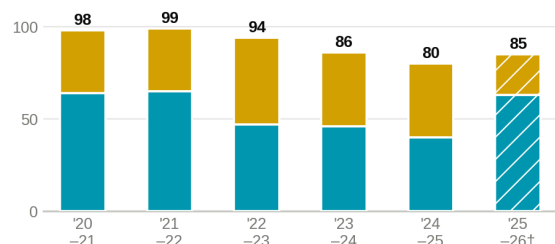


— District 112 — Oak Terrace — Red Oak — Northwood MS

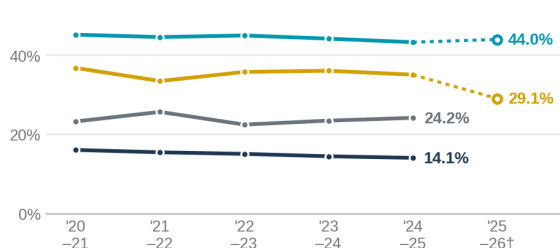
Enrollment change, indexed to 2020–21 = 100



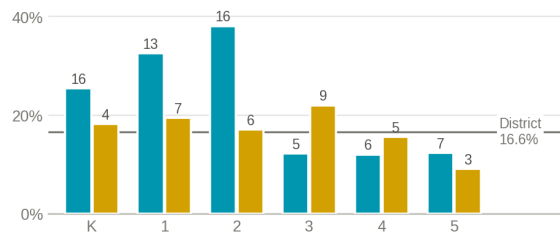
Kindergarten cohort, two program sites



English Learner share of enrollment



Students with IEPs by grade, 2025–26†



Sources: Illinois State Board of Education, Illinois Report Card 2020–21 through 2024–25, North Shore SD 112; District 112 classroom roster, April 2026.

Figure I.1. North Shore School District 112 enrollment and demographic trends, 2020–21 through 2025–26.

Sources: Illinois State Board of Education, Illinois Report Card 2020–21 through 2024–25; District 112 classroom roster, April 2026.

The District 112 Dual Language Program is a Spanish–English two-way immersion program that has operated for more than twenty-five years, available to students from kindergarten through eighth grade. The intended classroom ratio for the program is 50 percent native Spanish speakers or bilingual students and 50 percent native English speakers, with some students who speak another language at home. Instruction is delivered by bilingual-certified teachers.

The district describes the program as following a 50/50 language allocation model. Reading and writing are introduced in both English and Spanish beginning in kindergarten. In grades K–5, mathematics is taught primarily in English with connections to Spanish, and science and social studies alternate daily and are taught in Spanish. This description is the district's own, taken from its Request for Proposals. The relationship between the model as described and the model as delivered, including the analysis of the 2025–26 master schedule, is reported in Curriculum and Pacing Analysis section of this report (V.6), and the Program Model Analysis and in the Program Models Recommendations (VI.4) that follow.

The K–5 program currently operates at Oak Terrace and Red Oak Schools and will consolidate at Oak Terrace School beginning in 2027–2028. The middle school program is housed at Northwood Middle School. Students are encouraged to remain in the program from kindergarten through eighth grade in order to preserve the language balance the model depends on. Program graduates continue into Township High School District 113, where they may earn the Illinois Seal of Biliteracy and the Highland Park High School Dual Language Seal.

The district states the program's goals as high proficiency in both languages, high academic achievement, and positive attitudes toward language and diverse cultures.

The district is consolidating its dual language program from two schools into one as voted by the SD 112 Board of Education on Thursday, June 26, 2025. Therefore the findings that follow should be read as observations of a system under transition, and readers should keep in mind that the decline in English learner enrollment reported in the demographic data, is influenced by housing economics in the communities the program serves, not by the program's design, its quality, or its recruitment effort.

Scope: Grade Levels Under Evaluation

Although the dual language program at SD 112 is run from K-8, this evaluation addresses the K–5 dual language program as implemented at Oak Terrace and Red Oak Schools only.

Evaluation Design

The evaluation is grounded in the Guiding Principles for Dual Language Education, Third Edition (Center for Applied Linguistics, 2018), and in the Three Pillars of Dual Language Education: high academic achievement, bilingualism and biliteracy, and sociocultural competence. The three pillars are named directly in the district's RFP as the framework against which program effectiveness should be assessed, so the evaluation framework matches the district's own stated criteria.

Six Evaluation Questions organize the findings, each mapped to multiple strands of the Guiding Principles and examined across the three pillars. EQ 1 (English language proficiency), EQ 2 (Spanish language proficiency), EQ 3 (academic achievement), and EQ 4 (biliteracy assessment), EQ 5 (class size and composition), and EQ 6 (curriculum and pacing) A seventh overarching question, EQ 7, addresses the program model and is analyzed within the instructional model deliverable rather than as a findings chapter.

One of the seven strands of practice that organize this study, the Family and Community Strand, was not made a primary focus of this evaluation as the district has already been prioritizing family perspectives and collecting data through its own Dual Language Task Force. The district has surveyed families PreK-5 at both schools and K–5 dual language staff in the 2025–26 school year, and has gathered information through the standing work of the Bilingual Parent Advisory Committee. The decision to limit the coverage of this report around Family and Community Strand was not a judgment about the importance of the domain, and the implications for the report are set out in the Limitations of the Evaluation Study section (VIII.4).

The full evaluation question framework, including the composite EQ-by-pillar matrix, is presented in Section IV - Overall Methodology & Data Sources. Individual Evaluation Question Maps appear in Appendix A.

Evaluation Questions

1. To what extent are English learners (ELs) developing English language proficiency, and progressing toward reclassification? (RFP Sec. C)
2. To what extent are all students developing Spanish language proficiency? (RFP Sec. C, D)
3. What is the achievement of EL students in the Dual Language program for Language Arts and Mathematics? (RFP Sec. C)
4. How might a biliteracy assessment framework be conceptualized for the dual language program in SD 112? (RFP Sec. C, D)
5. What are the optimal class sizes and student composition models for K–5 dual language classrooms in SD 112? (RFP Sec. B)

6. How coherent, vertically aligned, and consistently paced is the dual language curriculum across grade levels in SD 112? (RFP Sec. C, D)
7. **Overarching Question:** Is the currently targeted 50/50 model effective for SD 112, or should the district transition to an 80/20 model? (RFP Sec. C, D)

EQ 7: how effective the current 50/50 model is, and what model is recommended going forward is the culminating question. It does not have its own findings chapter and is analyzed in the instructional model deliverable (VI.4), which is the only place in the report where a program model is recommended.

Data Sources

The evaluation draws on the following sources:

Student achievement and language proficiency data. A minimum of three years of data including the Illinois Assessment of Readiness (ELA and mathematics), NWEA MAP in English and Spanish, ACCESS for English Learners, and Spanish language arts assessments. Analysis was planned as longitudinal by cohort and school site, and disaggregated by grade, and English Learner status.

Classroom observations. Structured observations of literacy and mathematics instruction across K–5 at Oak Terrace and Red Oak Schools, conducted April 22–30, 2026 by J. Hilliard, C. Sánchez-López, and M. Gottlieb, using a rubric aligned to research-based dual language practice and documenting language of instruction by subject and time. The proposed visit schedule records approximately 72 scheduled observation entries across roughly 46 distinct classrooms, 41 entries at Oak Terrace and 31 at Red Oak, with two entries marked not conducted owing to teacher absence. Most classrooms were observed twice, once in language arts and once in mathematics.

Teacher and specialist interviews and focus groups. Individual semi-structured interviews of 30–45 minutes and 60-minute grade-level focus groups with dual language teachers, special education teachers and specialists, ESL and bilingual specialists, and instructional coaches.

Administrator and instructional leader interviews. Individual semi-structured interviews and 60-minute focus groups with building principals at both sites, the Assistant Superintendent for Teaching and Learning, the Director of Multilingual Learners, curriculum directors, special education directors, and the incoming Superintendent.

Parent focus groups. Focus groups conducted separately in English and Spanish with dual language program parents.

Student focus groups and work samples. Grade-level focus groups of approximately 30 minutes, together with review of student work samples and portfolios, including Spanish writing samples scored against district Spanish Language Arts rubrics and WIDA Marco DALE proficiency level descriptors.

Document review. Program goals and historical documentation; curriculum maps, pacing guides, and instructional materials; assessment data and bilingual assessment protocols spanning the pre- and post-COVID periods; teacher rosters, professional development records; policies on student placement, language allocation, and program transitions; and de-identified class language rosters and reclassification data.

Detailed methods, including sampling rationale, instrument development, and analytic approach, are presented in Section IV, Overall Methodology and Data Sources, which describes the methods for all six evaluation questions in one place.

The Evaluation Team

The evaluation was conducted by a team of four. John F. Hilliard, MEd, President of Paridad Education Consulting, served as Project Director, with responsibility for instructional model analysis, curriculum framework development, classroom observations, administrator interviews, student focus groups, data organization, and the class size analysis. Cristina Sánchez-López, MEd, Co-President of Paridad, served as Co-Principal Investigator and On-Site Coordinator, leading the English Learner needs analysis, classroom

observations, teacher interviews, family engagement in English and Spanish, sociocultural competence assessment, and special education considerations. Margo Gottlieb, PhD, served as Senior Evaluation Consultant, responsible for the biliteracy assessment framework, quantitative analysis, and student achievement reporting. Fred Genesee, PhD, Professor Emeritus at McGill University, served as Dual Language Researcher and Data Analyst, providing research design consultation and interpretation of bilingualism and biliteracy outcomes.

Exceeding Original Scope

In two respects the evaluation was executed at a scale beyond what the original agreement required.

First, the agreement contemplated classroom observations drawn from a sample of dual language classrooms with a minimum of 2 classroom observations per grade level per school. Instead, every dual language section at Oak Terrace and Red Oak, 26 of 26, was scheduled for observation, and a majority of classrooms were visited on two separate occasions rather than one, in most cases across different instructional blocks. The observational findings reported in Analysis Section (IV) consequently describe the program as a whole.

Second, the agreement called for an inventory of the biliteracy assessment instruments in place. The team instead documented the district's full K–5 assessment system, a total of 91 discrete components spanning eight adopted curricular programs together with the district, state, and local measures administered alongside them. That wider frame is what permits the analysis in Biliteracy Assessment Analysis (V.4) to distinguish measures that are genuinely absent from measures that have been deliberately replaced, and to identify assessments the district already has access to but does not currently administer. Both expansions were undertaken because the narrower scope would have yielded incomplete findings.

Organization of the Report

The report is organized in eight numbered sections. Sections I through III establish the foundation: this introduction, the research base for dual language education, and the design of the evaluation. Section IV states the overall methodology and data sources once, up front. Section V reports the data analyses and findings by evaluation question. Section VI presents the four contracted deliverables and all recommendations. Section VII sets out the implications of the evaluation. And finally, Section VIII states the limitations of the study.

The division between Evaluation Question Data Analysis & Findings Section V and Recommendations Deliverables Section VI is deliberate. The Section V reports evidence and includes no recommendations; each finding is assigned an identifier (V.1-F1, V.2-F1, and so on) so that it can be traced forward into the Recommendation Section VI. Section VI is where every recommendation in the report appears once, and each deliverable chapter cites the findings it builds on. Findings are in Section V and recommendations are in Section VI; a reader who wants to test whether a recommendation is grounded follows the finding identifiers between them.

Section V comprises six chapters, one per evaluation question: V.1 English Language Proficiency, V.2 Spanish Language Proficiency, V.3 Academic Achievement, V.4 Biliteracy Assessment, V.5 Class Size and Composition, and V.6 Curriculum and Pacing. Section VI comprises four deliverable chapters in the following order: VI.1 Class Size Recommendations VI.2 Curriculum Alignment and Pacing Recommendations VI.3 Biliteracy Assessment Framework Recommendations, and VI.4 Program Model Recommendations.

EQ 7, the program-model question, does not have its own chapter in Section V and is analyzed in Program Models Recommendation (VI.4), the only place in the report where a program model is recommended, and it draws on all six findings chapters.

SD 112 DUAL LANGUAGE PROGRAM EVALUATION · FINAL REPORT

Report Structure at a Glance

Eight numbered sections · Six evaluation questions · Four deliverables

FRONT MATTER

Cover Page · Table of Contents · List of Figures and Tables · Executive Summary

The Executive Summary previews the program's strengths, what the evaluation found, and the recommendations under each of the four deliverables.

FOUNDATION & METHOD

WHY THE STUDY EXISTS · WHAT IS KNOWN · HOW IT WAS DESIGNED · HOW IT WAS DONE

SECTION I

Introduction & Context

Purpose and scope, district and program context, the evaluation design and questions, data sources, the evaluation team, and how the report is organized.

SECTION II

Research Foundations

The research base for dual language education, from the Prism Model and the three pillars to biliteracy, assessment, and program structure, and what it means for SD 112.

SECTION III

Design for Dual Language Evaluation

The four RFP deliverables, the seven evaluation questions that lead to them, and how each deliverable maps to its questions.

SECTION IV

Overall Methodology & Data Sources

A census rather than a sample; data sources, and what was not collected; the April 2026 field week; analytic methods; reporting conventions. Stated once, so Section V stays on findings.

SECTION V · DATA ANALYSES, REPORTING & FINDINGS (EQ 1–6)

FINDINGS STAND ALONE — NO RECOMMENDATIONS HERE

V.1

EQ 1 · STRAND 4

English Language Proficiency

ACCESS growth, spring 2023–2025, and reclassification against the Illinois 4.8 criterion.

feeds VI.3 · VI.4

V.2

EQ 2 · STRAND 4

Spanish Language Proficiency

Grade 4–5 Spanish writing samples, classroom observation, curriculum review, and focus groups.

feeds VI.3 · VI.4

V.3

EQ 3 · STRAND 4

Academic Achievement

IAR annual and NWEA MAP interim measures in language arts and mathematics, three years.

feeds VI.3 · VI.4

V.4

EQ 4 · STRAND 4

Biliteracy Assessment Framework

The measures in place for literacy and oral language in English, in Spanish, and across both.

feeds VI.3 · VI.4

V.5

EQ 5 · STRAND 1

Class Size and Composition

Current K–5 class sizes and speaker-composition ratios, and the configurations that would strengthen the two-way model.

feeds VI.1 · VI.4

V.6

EQ 6 · STRAND 2

Curriculum Coherence and Pacing

Scope and sequence, pacing, materials, and language allocation across grades K–5.

feeds VI.2 · VI.4

findings carry forward by ID — V.1-F1, V.2-F1 ... — and every recommendation in Section VI cites the findings it rests on

SECTION VI · DELIVERABLES AND RECOMMENDATIONS (RFP SEC. D)

EVERY RECOMMENDATION APPEARS HERE, ONCE

VI.1

DELIVERABLE 1

Class Size Recommendations

Class sizes and speaker-composition ratios for K–5 dual language classrooms.

anchored V.5

VI.2

DELIVERABLE 2

Curriculum Alignment, Coherence & Pacing Recommendations

A unified K–5 scope, sequence, pacing, and language allocation, built on a proposed biliteracy curricular framework.

anchored V.6

VI.3

DELIVERABLE 3

Biliteracy Assessment Framework Recommendations

A balanced biliteracy assessment system measuring literacy and oracy in both languages, K–5.

anchored V.4 · also V.1, V.2, V.3

VI.4

DELIVERABLE 4

Program Model Recommendations

How effective the current 50/50 model is, and what program model is recommended going forward.

fed by V.1 – V.6

EQ 7 — the program-model question — is not a chapter in Section V. It is analyzed inside VI.4, the only place a program model is recommended, and draws on all six findings chapters.

SECTION VII

Prioritized Recommendations for Program Implementation

The strengths the program can build on; the conditions for continued success in language balance, enrollment planning, instruction, and measuring biliteracy; and the Section VI recommendations sequenced into short-, middle-, and long-term actions. **It cites recommendations by ID and adds none of its own.**

SECTION VIII

Limitations of the Evaluation Study

What the design can and cannot support, stated once: the absence of a Spanish language proficiency measure, a single field week of classroom observation, and the line between evidence and professional judgment.

GLOSSARY

Dual language and assessment terms, and the assessments named in the report.

REFERENCES

Sources cited across the report: the research base and district documents.

APPENDICES

A Evaluation Question Maps / Guiding Principles crosswalk · **B** Data collection instruments · **C** Interview and focus group protocols · **D** Proposed year-long biliteracy assessment schedule

CLOSING MATTER

About the Authors · Acknowledgments

Figure I.2. Report structure at a glance.

II. RESEARCH FOUNDATIONS

Purpose and Approach

This section sets out the research and theory behind the evaluation of School District 112's K–5 dual language (DL) program. It synthesizes peer-reviewed research, large-scale longitudinal studies, and professional guidance documents on two-way dual language education, with particular attention to studies of 50% English/50% Spanish program models comparable to SD 112's design. The review is intended to support professional judgment about SD 112's specific context. It is organized around the same Three Pillars framework the district has adopted to guide program design and evaluation, and each subsection is cross-referenced to the evaluation question(s) it most directly informs (EQ1–EQ7).

Theoretical Framework: The Prism Model and the Three Pillars of Dual Language Education

The conceptual foundation for this evaluation is based on two related bodies of work. Thomas and Collier's multi-decade longitudinal research program, most recently validated in Collier and Thomas (2017) across 36 districts, 16 states, and more than 7.5 million student records, established the Prism Model, which holds that language development, academic achievement, cognitive development, and sociocultural processes are interdependent and must be supported simultaneously, in both languages, for language-minority students to reach parity with their peers. This research found that well-implemented, long-term dual language programs close the achievement gap between English learners (ELs) and native English speakers within five to six years, compared with roughly half the gap closed by short-term transitional or English-only models (Thomas & Collier, 2002; Collier & Thomas, 2017).

This work is the direct antecedent of the Three Pillars framework articulated in Guiding Principles for Dual Language Education (3rd ed.) (bilingualism and biliteracy, academic achievement, and sociocultural competence) delivered through seven strands of practice: program structure, curriculum, instruction, assessment and accountability, staff quality and professional development, family and community, and support and resources (Howard et al., 2018). Paridad SD 112 has adopted this framework to organize its evaluation questions, and it is used throughout this report as the framework for interpreting findings against the research base.

English and Spanish Proficiency Development

A consistent finding across the research base is that the allocation of instructional time between English and the partner language does not translate in a simple, linear way into English proficiency outcomes. Lindholm-Leary's studies of California DL programs found that 50/50 and 90/10 models produced statistically similar rates of English proficiency by grade 8 (87% vs. 89% of Latino ELs reaching English-proficient status) despite the 90/10 model allocating far less early instructional time to English, and that stronger partner-language proficiency at program entry predicted higher subsequent English proficiency (Lindholm-Leary, 2012a, 2019).

Umansky and Reardon (2014) similarly found that DL students reclassify more slowly, on average, than students in English-immersion settings, but that the early "reclassification advantage" for English immersion reverses by high school. This is evidence that reclassification speed alone is not an adequate measure of program effectiveness.

Large lottery-based studies of Portland's DL program corroborate this pattern for English-partner-language development: by grade 8, DL-assigned EL students reach intermediate English-partner-language proficiency benchmarks well above comparison peers (Steele et al., 2017; RAND Corporation, 2017). Fortune (2015) provides one of the few empirical benchmarks

for oral language proficiency growth in Spanish immersion settings using CAL's SOPA/COPE instruments, finding significant, stepwise gains in fluency, grammar, vocabulary, and listening comprehension from kindergarten through grade 8 but cautions that most students reach only Novice High–Intermediate Low proficiency by grade 8, short of full bilingual fluency. This gives SD 112 a realistic reference point for setting benchmarks.

The National Academies' synthesis of the K–12 English-proficiency literature (NASEM, 2017, Ch. 6) adds an important calibration point for SD 112's ELP benchmarks (EQ1): across dozens of U.S. studies using widely varying reclassification instruments, the most consistently reported time to English reclassification is five to seven years, with several large studies estimating a median of eight years or more (e.g., Umansky & Reardon, 2014) and a substantial share of ELs, commonly (20–45%) do not reclassify even after five years of schooling. Progress is not linear: reclassification is fastest in the first year or two after school entry and slows markedly as students move from intermediate to higher English proficiency levels, a pattern Saunders and O'Brien's (2006) synthesis found to be "strikingly consistent" across dual language, transitional, and English-only program types alike. For SD 112, this means that a plateau in ELP growth in the upper-elementary grades is not, by itself, evidence that the 50/50 model is underperforming, because the same slowdown appears in the research regardless of program model. NASEM (2017) also found that students who enter kindergarten with stronger academic proficiency in both English and the home/partner language reclassify at meaningfully higher rates (e.g., Thompson, 2015, found a 24-percentage-point advantage), reinforcing Lindholm-Leary's finding that partner-language strength supports English development.

The same synthesis is also a caution for EQ2 (Spanish Language Proficiency). Even within well-implemented dual language programs, children's home-language proficiency needs ongoing support. NASEM (2017, Conclusion 6-5) concludes that ELs are at risk of losing ground in the partner language when English exposure begins early and intensifies, and that this risk is present "even when students are in dual language programs." Loss is moderated by continued, substantive use of the partner language at school and at home, and is more pronounced the more English-dominant a program's actual (as opposed to designed) allocation becomes, a finding SD 112 should weigh directly against the allocation-fidelity gap documented elsewhere in this evaluation (see EQ6; Section 2.6 below).

Academic Achievement Outcomes

The strongest and most consistent finding in the literature is that two-way DL programs do not lower, and typically raise, academic achievement for both ELs and native English speakers. Thomas and Collier (2002) found two-way programs produced the strongest long-term outcomes of any model studied nationally, with former-EL students reaching or exceeding grade-level English reading benchmarks by grades 3–5. Lottery-based causal studies (the strongest available research design, since they address self-selection into DL programs) corroborate this at scale: Steele and colleagues (2017, 2018) found DL-assigned students in Portland outperformed peers on state reading assessments by the equivalent of seven to nine months of additional learning by grades 5 and 8, with effects similar across native-language background and one-way/two-way models. Bibler (2021) found comparable, though more modest, positive effects (approximately 0.04–0.05 SD per year) in Charlotte-Mecklenburg schools.

Several studies suggest the DL advantage emerges gradually. Valentino and Reardon (2015) found DL students can trail peers in early grades before surpassing them by grade 7. Math achievement evidence is more mixed: Steele and colleagues (2015, 2017) found no significant DL math effect in Portland, while Lindholm-Leary and Borsato's (2005) longitudinal follow-up of Hispanic high schoolers from two-way elementary programs found native Spanish speakers moved from below grade level in grade 2 to above grade level by grade 6 in math, sustaining college-preparatory math enrollment into high school.

The research draws a clear distinction, relevant to how SD 112 should interpret its own IAR and NWEA data, between social/conversational English and academic English, the register of language used to acquire content knowledge, argue, compare, hypothesize, and synthesize (Chamot & O'Malley, 1994; Goldenberg & Coleman, 2010, as reviewed in NASEM, 2017). Academic-language proficiency develops more slowly than social language and is what standardized content assessments in English actually demand of ELs, regardless of program model. NASEM (2017, Conclusion 6-1) concludes that ELs typically need the full five-to-seven-year window (sometimes longer) to reach the academic-English proficiency that content-area assessments assume, which means a Grade 3–5 EL who is still developing academic English is not, on that basis alone, an indicator of program or instructional failure. The same lag would be expected in any program model.

Biliteracy Development and Assessment

Compared with English-language outcomes, the research base on partner-language (*Languages Other Than English - LOTE) literacy assessment is thinner and more field-based than experimental. This is a gap that SD 112 should recognize as common across the field and not a sign of local shortcomings, as it builds a biliteracy assessment framework. Howard and Sugarman (2009) note that the simultaneous-biliteracy approach to initial literacy instruction used in 50/50 models, the design SD 112 has adopted, has less direct empirical study than the sequential, partner-language-first approach typical of 90/10 programs, and flag cross-linguistic coordination between English and Spanish literacy instruction as a key element to monitor.

Heineke and Davin's (2025) study of Chicago's Seal of Biliteracy implementation offers a useful applied model: districts set interim, grade-level biliteracy benchmarks by "planning backward" from a high-school proficiency target, using instruments such as STAMP and AAPPL. Mihaly, Arellano, and Prier (2022) further show that, in a large New Mexico district, students who earned biliteracy recognition were more likely to enroll in four-year colleges, and that former ELs were not disadvantaged relative to English-proficient peers in earning it though many students met partial requirements without completing one. SD 112 should watch for similar attrition across grades K–5. Lindholm-Leary (2012b) argues that because English and partner-language reading achievement are highly correlated, accountability systems that assess English literacy only will systematically undercount DL students' full academic progress.

Basic-research evidence on cross-linguistic transfer strengthens the case for a biliteracy assessment framework that gives equal weight to both languages. NASEM (2017, Conclusions 6-3 and 6-4) concludes that the two languages of bilingual children do not develop in isolation: skills and strategies acquired in one language, especially those related to literacy, transfer to and support the other, and this transfer is strongest between typologically similar languages and in domains tied to literacy (decoding, comprehension strategies, morphosyntactic awareness). In practice, a biliteracy assessment framework that treats Spanish literacy purely as a parallel, independently-progressing track misses evidence that growth in one language can be a leading indicator for the other, an argument for building cross-linguistic progress monitoring into EQ4's instrumentation rather than scoring English and Spanish literacy as fully separate constructs.

Program Structure: Model Design, Class Size, and Curricular Coherence

Research comparing 50/50 and 90/10 program models generally finds outcomes converge by the upper elementary/middle grades, though 90/10 students often show an early partner-language advantage that narrows over time (Lindholm-Leary, 2004, 2012a). Howard, Sugarman, and Christian's (2003) synthesis of two decades of two-way immersion research identifies common conditions of successful implementation (early stakeholder engagement, a clear and consistently communicated program vision, and administrative and family support) that concern program structure and are therefore within SD 112's control regardless of which model variant it selects.

On class size specifically, the general-education literature offers limited support for assuming that class-size reduction alone will substantially change outcomes: Slavin's (1989) classic review found that even substantial reductions (for example, from roughly 31 to 16 students) produce only small, often non-durable achievement gains, and that effects of practical significance require reductions to near one-to-one tutoring ratios. For DL programs specifically, class composition, the ratio of native English speakers to partner-language speakers, is treated in the literature mainly as a way to protect the two-way model's language balance, reinforcing that SD 112's class-size question should be evaluated jointly with composition ratio, not class size alone. On curricular coherence, both the English Learners Success Forum's materials-evaluation guidelines and Thomas and Collier's (2002) findings show that materials and pacing must sustain cognitively demanding, grade-level content in both languages rather than simplifying content for developing bilinguals; simplification is associated with the weaker long-term outcomes seen in short-term transitional models.

Sociocultural Competence and Culturally Sustaining Pedagogy

The Three Pillars framework treats sociocultural competence as equal in importance to language and academic outcomes. Paris and Alim (2014) provide the conceptual basis for this pillar as it has evolved in the field, arguing that schools should aim not merely to accommodate students' home languages and cultures but to actively sustain and extend their linguistic and cultural pluralism as a matter of educational equity. This framing is reflected operationally throughout Guiding Principles for Dual Language Education (Howard et al., 2018), which embeds cultural affirmation, equitable status for both languages, and community partnership into every strand of practice rather than isolating it as a stand-alone program component.

Inclusive Dual Language Education for Bilingual English Learners (ELs) with Disabilities

Because SD 112's evaluation questions and deliverables cover the full range of learners the program serves, this evaluation also draws on the growing research base on EL bilingual learners with disabilities in dual language programs. This research is directly relevant to the Instructional Model Analysis (EQ7), the Biliteracy Assessment Framework (EQ4), and the Curriculum & Pacing Framework (EQ6), each of which must also work for students receiving special education services.

A persistent misconception (that bilingualism is an added burden for children with disabilities, likely to overwhelm or confuse them) is not supported by the evidence (Sánchez López, Young, Genesee, & Hilliard, 2022; Meadvin, Montes Pate, & Whitney, 2026, summarizing Padía, Cioè-Peña, & Phuong, 2024). Reviewing research on young language learners (ELs in the US) with developmental language disorder, autism, and Down syndrome, Sánchez López et al. (2022) conclude that these children can become bilingual and generally do so with the same profile of strengths and challenges as monolingual children with the same disability; bilingual exposure does not exacerbate the disability, provided the learning environment offers sufficient, consistent exposure to both languages. At the school-age level, U.S. studies comparing EL students with disabilities to their peers with comparable disabilities in monolingual programs, Howard (2003), Lindholm-Leary (2005), Myers (2009), and Thomas and Collier (2014), found no significant achievement disadvantage for Dual Language participation, and in some cases EL students with disabilities outperformed their monolingual-program counterparts in reading and math despite less overall exposure to English; The SEAL brief's analysis of California data found that students who remain long-term ELs are disproportionately likely to carry a disability designation, 28% of long-term ELs statewide, versus 11% of California students overall (Meadvin, Montes Pate, & Whitney, 2026). The same brief and the CAL research review point to practical implications for the Biliteracy Assessment Framework and Curriculum & Pacing Framework deliverables: sustained collaboration among bilingual/DL teachers, special education staff, and

English-language-development specialists, rather than siloed service delivery, is associated in both sources with more coherent support for this population. Sánchez López et al. (2022) further note that maintaining and developing the home language for ELs, rather than narrowing to English, has been shown even in children with developmental language disorders to support English acquisition, the same cross-linguistic-transfer logic already noted in Section 2.5 for the biliteracy assessment framework generally.

Program Effectiveness, and Fidelity

Because DL enrollment is typically voluntary, a persistent methodological caution across this literature (flagged explicitly by Howard, Sugarman, and Christian (2003) and by Valentino and Reardon (2015)) is self-selection: families who choose DL programs may differ from those who do not in ways that confound simple outcome comparisons. The strongest recent evidence addresses this directly through lottery-based designs, which approximate random assignment among applicants.

Steele and colleagues (2018) found that roughly one-third of DL's positive effect on ELA achievement in Portland was associated with differences in per-pupil expenditure, indicating that resources and classroom conditions account for a substantial share of the program's effect. The companion RAND cost analysis (Steele et al., 2015) found DL per-pupil costs scaled roughly proportionally with enrollment share and were not substantially out of line with mainstream instruction.

Synthesis: Research Base, Gaps, and Implications for SD 112

Overall, the research strongly supports two-way dual language education as an effective model for both English learners and native English speakers, with the largest and most rigorous studies (lottery-based designs and multi-decade, multi-district longitudinal studies) showing consistent positive effects on English reading achievement and no evidence of academic harm from instructional time provided in two languages.

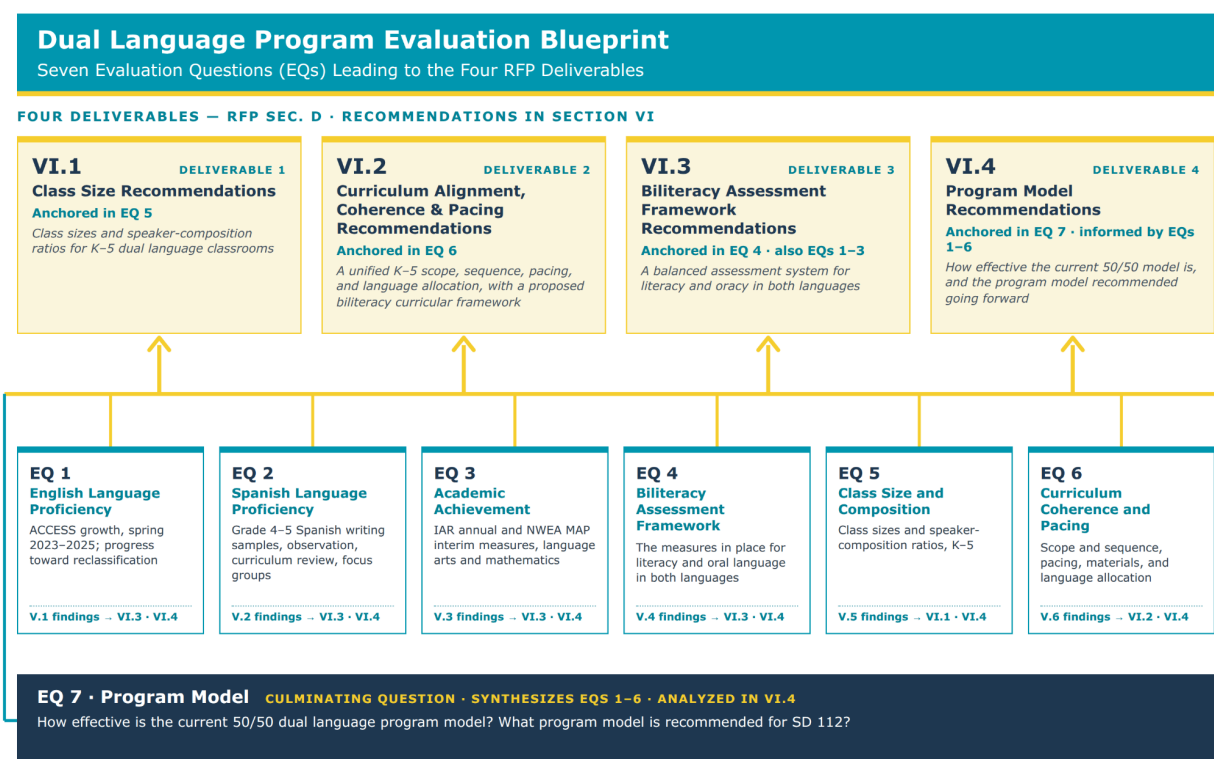
The evidence is strongest for English-language and general academic outcomes, and comparatively thinner in two areas central to this evaluation: simultaneous, 50/50-style biliteracy development specifically, and direct assessment of partner-language proficiency and literacy, where the field still relies more on practitioner-developed frameworks than experimental study. This gap exists across the field, not only in SD 112, and it makes the biliteracy assessment framework this evaluation is designed to help build (EQ4) all the more important. The findings presented in the sections that follow should be read against this base: where SD 112 data track the patterns established in this literature (for example, an early-grades achievement gap that narrows over time, or partner-language proficiency that lags but is not static) that pattern is consistent with a healthy, developing program rather than evidence of underperformance.

III. DESIGN FOR DUAL LANGUAGE EVALUATION

The Four Program Evaluation Deliverables

School District 112's RFP directs the evaluation to determine the effectiveness of its dual language education program according to the Three Pillars of Dual Language. The Pillars are high academic achievement, bilingualism and biliteracy, and sociocultural competence. Taking the three pillars of dual language education as foundational and non-negotiable, along with their accompanying Guiding Principles of Dual Language Education, this section presents the overall program evaluation design.

As illustrated in Figure III.1, the Dual Language Program Evaluation Blueprint spells out the evaluation design. It is a flowchart that delineates the relationship between the evaluation's four deliverables, shown as rectangles at the top of the chart, and the evaluation questions (EQs) from which they are drawn. Six of the seven EQs feed into the four specified deliverables described in IIC. The 7th EQ at the bottom of the chart that seeks to determine the overall effectiveness of the current 50-50 instructional model, is overarching with broad coverage across the set of EQs and deliverables.



Paridad Education Consulting · SD 112 Dual Language Program Evaluation

Figure version September 28, 2026

Figure III.1. The Dual Language Program Evaluation Blueprint, the seven EQs leading to the four RFP deliverables

The Seven Evaluation Questions (EQs) Leading to the Deliverables

The Dual Language Program Evaluation Blueprint revolves around seven evaluation questions (EQs 1-7) that represent a synthesis of the information and requests within School District 112's RFP; namely, a.) the four deliverables, b.) the three dual language education pillars, and c.) the seven guiding principles for dual language education. The EQs organize the entire dual language program evaluation. Together the EQs capture the quantitative and qualitative data that support the evidence for each deliverable.

As delineated below, each EQ consists of four parts. The bolded question at the top guides the evaluation, followed by the section(s) of the RFP from which it is drawn. It is then restated *in operational terms* in finer detail with specific reference to how it is presented in the report. The *Primary focus* states the main data source(s) or most salient aspect of the EQ. The last part of the EQ names the most applicable *Deliverable*.

EQ 1	To what extent are students developing English language proficiency, and are English Learners progressing toward reclassification? (RFP Sec. C) In operational terms: <i>How much English language proficiency did English Learners gain each year from spring 2023 to spring 2025, measured on the WIDA ACCESS for ELLs composite scale score, and how many students were reclassified out of English Learner status over the same span against the Illinois 4.8 criterion?</i> Primary focus: ACCESS for ELLs™ data and reclassification rates for ELs. Deliverable: Instructional Model Analysis & Recommendations
EQ 2	To what extent are all students developing Spanish language proficiency? (RFP Sec. C, D) In operational terms: <i>What do grade 4–5 Spanish writing samples, classroom observations, curriculum review, and teacher, parent, and student focus groups show about students' Spanish proficiency and about the conditions producing it?</i> Primary focus: Assessment of Spanish language literacy Deliverable: Recommendations for a Comprehensive Bilingual Framework
EQ 3	What is the achievement of EL students in the Dual Language program for Language Arts and Mathematics? (RFP Sec. C) In operational terms: <i>What do the data show about achievement in the dual language program on the annual state test, and the interim measures for mathematics and language arts separately, English and Spanish separately, across the three years for which data exist?</i> Primary focus: State and district Language Arts and Mathematics annual and interim standardized achievement tests Deliverable: Instructional Model Analysis & Recommendations
EQ 4	How might a comprehensive bilingual assessment framework be conceptualized for the dual language program in SD 112? (RFP Sec. C, D) In operational terms: <i>What data and what instrumentation are in place to capture dual language students' literacy and oral language (listening and speaking) development in English, in Spanish, and across both languages?</i> Primary focus: How might SD 112 plan for a bilingual trajectory with crosslinguistic connections to promote language and literacy development? Deliverable: Recommendations for a Comprehensive Bilingual Framework
EQ 5	What are the optimal class sizes and student composition ratios for K–5 dual language classrooms in SD 112? (RFP Sec. B) In operational terms: <i>Do the class sizes and composition ratios currently in place support the two-way model, and what configurations does the evidence indicate would strengthen it?</i> Primary focus: Program Structure, class size and speaker-composition ratios as the structural conditions that preserve the two-way model Deliverable: Recommendations for Dual Language Class Size Analysis
EQ 6	How coherent, vertically aligned, and consistently paced is the dual language curriculum across grade levels in SD 112? (RFP Sec. C, D)

	In operational terms: <i>Do the scope and sequence, the <u>pacing calendars</u>, the materials, and the <u>language allocation</u> add up to one coordinated K–5 program or to two programs taught side by side?</i> Primary focus: Curriculum alignment across the <u>Three Pillars</u> Deliverable: Recommendations for Curriculum & Pacing Framework
EQ 7	How effective is the current 50/50 dual language <u>program model</u>? What <u>program model</u> is recommended for SD 112? In operational terms: <i>What does the combined evidence show, and how does it point to the most appropriate dual language <u>program model</u> for <u>English learners</u> in District 112?</i> Primary focus: To what extent are the <u>Three Pillars</u> operationalized across the current program structure? Deliverable: Instructional Model Analysis & Recommendations

Mapping the RFP Deliverables to The Evaluation Questions (EQs)

To summarize this section of the report, the Dual Language Program Evaluation Blueprint identifies the relationship between the four required Deliverables and how they manifest in and across the seven evaluation questions (EQs). The Blueprint rests on the three pillars of dual language education and the seven guiding principles.

Because each Deliverable is complex, there is not a one to one correspondence between the four sets of recommendations listed below and the EQs. Rather, to provide thorough responses, each Deliverable has been informed by a cluster of EQs.

1. **Instructional Model Analysis & Recommendation:** based on EQ 7 and informed by EQs 1–6, culminating in the recommendation to continue a 50/50 instructional model or transition to an 80/20 model
2. **Comprehensive Biliteracy Assessment Framework Recommendations:** drawn from EQs 2 and 4, with recommendations for how to design a comprehensive balanced biliteracy assessment system addressing both English and Spanish
3. **Curriculum & Pacing Framework Recommendations:** based on EQ 6, also informed by EQ 7, with recommendations for how to support K-5 curricular coherence across Spanish and English
4. **Dual Language Class Size Analysis and Recommendations:** based on EQ 5, with recommendations for class size and student composition in preparation for consolidating the dual language program at Oak Terrace school in the 2027-28 school year

Each Evaluation Question is repeated at the top of a page and then further detailed in a map (see Appendix A for the full set of 7 EQ maps) ; Figure III.2 is a display of the Sample Evaluation Question Map for the overall Program Model, (EQ 7) the most comprehensive of the EQs. The purpose of the map is to provide an overview of the primary targets of dual language education represented by the intersection of the three pillars of dual language and the seven strands of practice from the Guiding Principles. The resulting matrix of 21 cells defines the key aspects of dual language education. The color coding specifies the relative importance of each cell in contributing to the designated Deliverable- orange delineating the primary contributor, mustard, the secondary contributor, and light yellow, the least of the contributors.

Underneath the EQ matrix is the specific Deliverable that is addressed and a display of all the applicable data sources used to answer the EQ. The Quantitative sources are on the left and Qualitative data sources on the right, providing the evidence for the stated Deliverable.

Sample Evaluation Question Map - EQ 7 Program Model

How effective is the current 50/50 dual language program model? What program model is recommended for SD 112?

In operational terms: *What does the combined evidence show, and how does it point to the most appropriate dual language program model for English learners in District 112?*

Primary focus: To what extent are the Three Pillars operationalized across the current program structure?

Deliverable: Instructional Model Analysis & Recommendations

STRANDS of Practice (the means) advance → the pillars	Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Language allocation protects partner-language time	Articulated PreK–12 grade-level pathway	Equal status for both languages by design
2 Curriculum	Biliteracy trajectory; cross-linguistic bridge	Biliteracy trajectory; cross-linguistic bridge	Culturally sustaining, multicultural content
3 Instruction	Strategies for transfer & target-language output	Scaffolded, cognitively rigorous content	Affirms students' languages & identities
4 Assessment & Accountability	Assessment of proficiency & literacy in both languages	Documentation of grade-level progress; guides instruction	Indicators of cross-cultural growth; equitable measures across languages
5 Staff Quality & Prof. Development	Biliterate, dual-language trained teachers	Two-language content & pedagogy expertise	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures
7 Support & Resources	Time & materials for the <u>partner language</u>	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 7 ↓

DELIVERABLES

Instructional Model Analysis & Recommendations

DATA SOURCES · EQ 7

QUANTITATIVE	QUALITATIVE
– Language-allocation analysis	– Classroom observations (allocation)
– Proficiency & achievement trends	– Administrator interviews
– <u>Reclassification</u> trajectories	– Teacher interviews & focus groups
	– Document review (allocation plan)

Figure III.2. Anatomy of a Map of an Evaluation Question: Coverage of the Strands of Practice in the Guiding Questions for Dual Language Education in Relation to the Three Pillars of Dual Language Education along with its Data Sources

IV. OVERALL METHODOLOGY & DATA SOURCES

This section states how the evaluation was conducted. Method is not restated inside the evaluation-question chapters; each of those carries only the handling particular to its own data, and its own data limitations. Study-wide limitations are treated separately in Section VIII, and the misreadings that a particular set of recommendations invites are treated in the deliverable chapters at VI.N.7. This section covers only the first of these.

This section should be read alongside Section III, which sets out the evaluative framework and the six evaluation questions, and Section I.3, which describes the district conditions under which the fieldwork was carried out. Those conditions bear directly on method: implementation was observed during a transitional year, and the design decisions recorded below were made with that in view.

Evaluation approach and evaluator stance

The evaluation is a mixed-methods, framework-informed program evaluation. The evaluative criteria were designed and drawn from the Guiding Principles for Dual Language Education (Howard et al., 3rd edition, 2018), whose seven strands of practice and three pillars organize both the questions and the reporting. It is mixed-methods in that quantitative and qualitative evidence were collected and analyzed on their own terms and then combined when findings were drawn. Where the two converge, the report says so and rates the finding accordingly. Where they diverge, the report presents both readings and names the evidence that would settle the matter.

The design is utilization-focused. Each evaluation question was framed so that its answer feeds a decision the district has to make, and every question routes to one of four deliverables: the class size analysis, the curriculum and pacing framework, the biliteracy assessment framework, and the instructional model analysis. Questions that could be asked but would inform no pending decision were not pursued.

Census framing rather than sampling

The dual language program is small enough to study in full, and this evaluation did so. In the 2024–25 roster year the program comprised 26 sections and 492 students across two schools: 15 sections and 293 students at one site, 11 sections and 199 students at the other. The evaluation did not draw a sample from that population. It worked from records covering all of it, and it scheduled classroom observation in every one of the 26 sections.

- **Descriptive claims cover the whole program.** When this report states a rate, a count, or a distribution for the program, it is reporting the population value, not an estimate of one. Because there is no sampling error, none is reported.
- **Statistical significance plays a limited role.** Because the evaluation covers the whole population, there is no sample-to-population inference to test. The report therefore leads with magnitudes, percentiles, and denominators, and uses significance language only where an external analysis supplies it, as in the district’s value-added growth reports, where the significance flag is part of the vendor’s own metric.
- **Findings apply only to SD 112.** A census of this program supports strong statements about it, but not about dual language programs elsewhere. Where the report reaches beyond D112, it does so through the research literature in Section II, and labels the finding External.

A scheduling census is not an evidentiary census

This distinction is easy to overlook. All 26 dual language classrooms appear on the field-week schedule, which is what makes classroom observation a census in the scheduling sense. The evidentiary yield is smaller: 17 observation records were completed, and 14 of those contain

instructional evidence. In the remaining three the teacher was absent, the class was not in instruction. Statements resting on observation are therefore made against a denominator of 17 or 14, never 26, and the applicable denominator is printed with every observation figure.

The same care applies on the quantitative side, where coverage varies by measure and by year: the interim assessment record supports three-year comparison at grades 3–5 but not below, because grade coverage expanded during the period. Cross-year analyses are restricted accordingly, and the restriction is stated in the chapter that uses them.

Quantitative data sources

All quantitative evidence in this evaluation is extant district or state data. No assessment was administered for the purposes of this study. The sources below are grouped by what they measure.

English language proficiency

- **ACCESS for ELLs (WIDA), 2022–23 through 2024–25.** Per-school overall results were assembled into a student-matched longitudinal file. The three-year matched cohort is $n = 119$. Composite, oral language, and literacy scale scores and proficiency levels were analyzed, together with reclassification indicators.
- **WIDA consortium growth benchmarks.** Annual growth was compared against published consortium growth values by grade cluster, which supplies an external reference point the district's own record cannot provide, and which resolves the vertical-scale objection that would otherwise limit year-over-year scale-score comparison.
- **Reclassification and time-to-reclassification records.** Partially in hand.

Academic achievement and growth

- **NWEA MAP, Reading and NWEA MAP, Mathematics (English).** Fall and spring administrations. Conditional is the primary metric and achievement percentile the secondary one; RIT scale scores are not used for cross-year comparison.
- **NWEA MAP Lectura.** The district's only Spanish-language achievement measure. Reported on user norms, which permits trend and conditional-growth statements but does not permit an English-minus-Spanish gap calculation; no such calculation appears in this report.
- **Illinois Assessment of Readiness (IAR), English language arts and mathematics, grades 3–5.** Spring proficiency distributions and student growth percentiles.
- **FastBridge early literacy and oral reading measures (English).** Fall-to-spring screening data at K–2, used for early-literacy trajectory. Spanish forms exist in the same product line.
- **District value-added growth reports, 2022–23 through 2024–25.** Vendor-produced conditional growth summaries against an internal comparison group, including a dual language indicator. Read at band level only, never as an effect size.

Program structure, composition, and the assessment system itself

- **Class size and composition roster, 2024–25.** Section-level enrollment with English learner and IEP counts, reconciled to the program totals. Grade assignments were confirmed through four independent routes (the field-week schedule, the language-arts assessment files, arithmetic reconciliation to the program totals, and teachers' own identifications in the focus-group rounds) which closed the grade ambiguity present in the original roster file.
- **District assessment inventory, verified by the district.** A 91-component inventory of curriculum-embedded and district, state, and local assessments, returned with a use status for 89 components: 28 in use, 17 partial or at teacher discretion, 44 not used, 2 unanswered. It documents how the assessment system actually operates.
- **District data tracker.** The district's own status log of the data it collects and plans to collect. It is used here as evidence about the data system rather than as a data source, and it is the documentary basis for statements about what the district does not measure.

- **Enrollment, demographic, home-language, and attendance records.** District-level demographic context is taken from the Illinois Report Card for 2024–25.

What the quantitative record does not contain

Three points of data do not appear in the report and worth noting as their absence were not choices made by the evaluators.

- No assessment of mathematics achievement is administered grades 1-5. Kindergarten does include assessment of mathematics achievement in Spanish. Mathematics achievement in the dual language program is measured exclusively in English. There are end of unit assessments in Spanish in the areas of science and social studies that teachers administer and other assessments within the science and social studies curricula that are administered at teachers' discretion.
- There is no Spanish language proficiency instrument of record, and no Spanish writing samples were collected during the field week.
- No Seal of Biliteracy data were available.

These absences are reported as findings about the district's assessment system in Sections V.3 and V.4, and they constrain what this evaluation can conclude about the program's Spanish-language outcomes. That constraint is treated at VIII.1.

Qualitative data sources

All qualitative data were collected by the evaluation team during the April 2026 field week, with the exception of documents supplied by the district before and after it.

Administrator interviews

Six semi-structured interviews were conducted with ten administrators on April 21, 2026, at the district office, using the evaluation's administrator protocol. Participants spanned district leadership, teaching and learning, student services, and both building administrations. The corpus runs to approximately 43,000 words. Each interview was analyzed on its own and the six analyses were then cross-synthesized, with findings classified as unanimous, near-unanimous, single-source, or contested. Contested points were not resolved here; each was assigned to the evaluation question best able to address it.

Teacher focus groups

Thirteen teacher focus-group records were produced across the two sites using the evaluation's focus-group protocol. These records are typed field notes entered into the protocol's question tables. They are not transcripts, which limits how they can be used. They establish that a view was expressed and by roughly whom, and they do not establish how widely it was held, how strongly, or whether it was contested in the room. Prevalence claims are not made from them.

Student focus groups

Five student focus-group records were produced, at grades 3 and above only; no group was held below grade 3, and one site has no grade-4 group.

Classroom observation

Seventeen classroom observation records were completed during the field week, of which fourteen carry instructional evidence. Observations covered homeroom instruction in both languages together with specials, intervention, and accelerated learning settings.

Two versions of the observation instrument were in circulation, and their use coincided exactly with site. Every observation **at one site used a 36-item sequential form, and every observation at the other used a 21-item synthesized form** that omits fifteen items, among them curriculum, vertical alignment, pacing, standards alignment, language-allocation adherence, and biliteracy. No item-level comparison between the two sites is made anywhere in this report, and the confound is restated wherever observation evidence is used.

A further limitation affects how the observation evidence is analyzed. The completed forms were returned as scans in which the rating checkboxes are not machine-recoverable and, in the copies available for analysis, not reliably legible. The observation analysis therefore relies on the written evidence and field notes recorded on the forms, not on rubric scores. No observation finding in this report depends on a rating.

Document review

The evaluation reviewed the interview and observation protocols, curriculum and assessment documentation, building floor plans and room inventories, and the district's own data tracker. Several documents named by participants as governing practice (scope and sequence documents, pacing calendars, curriculum guides, professional development documentation,) were also reviewed.

What was not collected

Limited data were collected on the Dual Language Guiding Principles Family and Community strand. Focus groups for Spanish and English-speaking families were conducted. The strand is one of the seven that organize the evaluative framework, and the evaluation cannot address it. See VIII.4.

All primary data collection took place in a single eight-day field week, April 21–30, 2026. April 21 was given to administrator interviews at the district office. The remaining days ran on-site: two consecutive days at the first school, three days at the second, one split day across both, and a final day at the first. Each day was divided into eleven 30-not true.minute slots between 8:50 a.m. and 2:45 p.m., with a midday break.

The schedule was transcribed slot by slot into an 87-row log, which is the evaluation's record of coverage. It includes 34 homeroom observation blocks, 7 specials, 2 accelerated learning, 1 intervention, and 2 instructional coaching blocks, alongside 9 teacher focus-group blocks, 6 student focus-group blocks, 3 walk-throughs, and 11 unassigned slots.

What the week covered

- Every one of the 26 dual language classrooms appears on the schedule, 15 of 15 at one site and 11 of 11 at the other, so every classroom was scheduled for observation.
- The schedule deliberately avoided each grade's collaborative learning day at the first site, so no grade was observed on the day its instructional configuration was atypical. We consider this a strength of the design.
- Observation reached beyond homeroom instruction into specials, intervention, and accelerated learning, which is what makes it possible to say anything at all about the program's effective language allocation across the school day rather than only within the homeroom schedule.
- Finally, a single week provides only a limited snapshot of the program.

ANALYTIC METHODS

Quantitative analysis

- **Percentile metrics over scale scores.** Cross-year comparison uses achievement and conditional growth percentiles rather than scale scores, because scale meaning is not stable across years and grade coverage changed during the period.
- **Conditional growth percentile as the primary measure.** Conditional growth compares each student against academically similar starting points, so growth statements hold even though the program's composition differs from the comparison population.
- **Matched-cohort construction.** The English proficiency analysis is built on students matched across years, one row per student, so that year-over-year change is measured within students rather than across shifting cohorts. Unmatched records are excluded from cohort statements and included only in single-year descriptions.
- **External benchmarking where an internal comparison would mislead.** English proficiency growth is compared against published consortium growth values; the district's value-added reports are read at band level against their own internal comparison group and are never converted into effect sizes.
- **Spanish results reported as trend and conditional growth only.** Because the Spanish measure reports on user norms rather than nationally representative norms, its results support statements about direction and about growth relative to like starting points, and do not support comparing Spanish and English standing, so no such comparison is made.
- **Denominators printed, sites pooled.** Every figure in Section V uses its own denominator, and no figure disaggregates by site (see IV.7). Effect sizes do not appear in Section V.

Qualitative analysis

- **Coding to the framework.** Interview, focus-group, and observation evidence was coded to the seven strands of practice and three pillars that organize the evaluative framework, so that qualitative evidence is organized the same way as the evaluation questions.
- **Case-then-cross-case.** Each interview was analyzed as a case before any synthesis was attempted, which preserves the internal logic of each account and ensures that agreement across interviews is observed, not produced by reading them together.
- **Observation analyzed through coverage and evidence, not scores.** An item-coverage matrix records, for every rubric item, how many of the seventeen records contain evidence on it; the substantive analysis draws on a logged set of written evidence entries. Items with zero coverage are also reported, because a rubric item that no observation addressed tells us something about the instrument and about what the evaluation can answer.

Triangulation of the Data and What it Means

Findings are rated on the strength of their supporting evidence, using four ratings applied consistently across Section V.

WHAT IT MEANS	TYPICAL BASIS
Two or more independent sources agree, or a documentary record settles it	Assessment records plus district documentation; a count reconciled two ways
One source, or a source whose reliability is bounded	A single interview account; a focus-group statement without corroboration

WHAT IT MEANS	TYPICAL BASIS
Depends on research literature or regulation rather than on district data	A regulatory ceiling; a published research finding
Sources conflict, or the evidence needed does not yet exist	A contested account with the adjudicating data outstanding

Where quantitative and qualitative evidence point in different directions, the report states both and identifies the evidence that would settle the question, and does not automatically favor one form of evidence over the other.

De-identification and reporting conventions

Students

- Student-level files remain in the analytic workspace. They are not circulated, not attached to deliverables, and not placed in shared project spaces, consistent with FERPA obligations.
- No result is reported at a level of disaggregation that would identify an individual student. The program's cells are small: most grades run two sections per site and one grade at one site runs a single section, so a grade-by-site result is in some cases a single classroom.
- For that reason no figure in Section V disaggregates by site. This was a deliberate reporting choice; site-level data exist. The choice is noted wherever readers might expect a site breakdown.

Staff

- No teacher, administrator, or staff member is named anywhere in the report, in any deliverable, or in any appendix.
- Qualitative evidence is attributed in the aggregate and in neutral terms ("an administrator," "teachers at both sites") never to a role specific enough to identify a person.
- Grade and site are never given together in an attribution, because that pairing identifies a classroom and in one case identifies a single teacher.
- Interview and focus-group material is paraphrased rather than quoted where a verbatim quotation would carry identifiable voice. A quote-risk register maintained with the interview analyses applies to every deliverable, and material recorded as off the record is excluded entirely.

Instruments and naming

- Assessments are named by their correct current names, and a measure the state has discontinued does not appear in any figure or recommendation even where participants referred to it.
- Where an instrument exists in the district's licensed suite but is not administered, it is reported as a gap rather than as an unused option, because each calls for a different response.
- Figures show measures in use. Non-adopted components are listed in the accompanying tables, where the analysis can cite them.

Report conventions

- Sections are numbered in Roman, subsections in Arabic: V.1–V.6 for the evaluation-question chapters; VI.1–VI.4 for the deliverable chapters. Findings carry chapter-scoped identifiers and recommendations carry deliverable-scoped ones, so that a recommendation can name the findings it builds on

Summary of data sources and instruments

Table IV.1 summarizes the evidence base. The full version, with instrument descriptions, administration windows, and record counts, is Appendix B.

SOURCE	INSTRUMENT OR RECORD	COVERAGE	EQ
English <u>language proficiency</u>	<u>ACCESS for ELLs</u> (WIDA)	2022–23 to 2024–25; matched <u>cohort</u> n = 119	1
English <u>language proficiency</u>	WIDA consortium growth benchmarks	By grade cluster, published values	1
English <u>language proficiency</u>	<u>Reclassification</u> records	Program-wide; DL vs non-DL breakout missing	1
Achievement, English	<u>NWEA MAP</u> , Reading;; Mathematics	Fall/spring, 3 years; grades 3–5 for cross-year work	3
Achievement, Spanish	<u>NWEA MAP</u> Lectura	Fall/spring, 3 years; user <u>norms</u>	2, 3, 4
Achievement, English	<u>IAR</u> , ELA and mathematics	Grades 3–5, spring 2025	3
Early literacy	<u>FastBridge</u> early reading and oral reading	K–2, English; Spanish administration unconfirmed	3, 4
Growth, external	District value-added growth summaries	3 years; DL flag scope to be confirmed	1, 2, 3
Program structure	Class size and composition roster	26 sections, 492 students, 2024–25	5
Assessment system	District assessment inventory, verified	91 components; 89 answered	4
Assessment system	District data tracker	Full data plan, all four domains	2, 3, 4
Context	Enrollment, demographic,	Requested; Report Card used in the interim	3, 5, 7
Qualitative	Administrator interviews	6 interviews, 10 administrators, ~43,000 words	All
Qualitative	Teacher focus groups	13 records, both sites	2, 5, 6, 7
Qualitative	Student focus groups	5 records, grades 3 and above	—
Qualitative	Classroom observation	17 records, 14 with instructional evidence	2, 6, 7

SOURCE	INSTRUMENT OR RECORD	COVERAGE	EQ
Qualitative	Field-week slot log	87 slots, 8 days, both sites	All
Documents	Task force report, protocols, floor plans	As supplied	All
Documents	Scope and sequence, pacing, PD, prior evaluation	Named by participants; not supplied	6, 7
Documents	Board record for Section I.3	Pull list prepared	—
Not collected	Family and community primary data	None	—
Not collected	Spanish writing samples; Spanish LP instrument	None	2

Table IV.1. Evidence base at a glance

The missing Spanish-language measures limit what this evaluation can conclude about the Spanish half of the program. This information is addressed in Section VIII.

V. DATA ANALYSES, REPORTING & FINDINGS

This section reports what the evaluation found, one chapter per evaluation question. Findings are labeled with an identifier of the form V.4-F1 (chapter, then finding number). This section reports findings only; it contains no recommendations. Qualitative evidence is reported in the aggregate throughout, without attribution to individuals or roles. No participant is identified by grade and site together.

Introduction & Organization

Every chapter in this section runs the same subsections, in this order:

THE EVALUATION QUESTION 1. States the evaluation question verbatim (the What), then it 2. operationalizes it (the How?) and finally it connects it to the underlying theoretical framework of the study 3. The Three Pillars and the Guiding Principles and Strands of Dual Language Education.

DATA SOURCES Every source consulted for the chapter, quantitative and qualitative, each marked as in hand, still outstanding, or non-existent. Listing a source here means the chapter drew on it; whether it answers the question is judged later.

METHODOLOGY The procedures run on those sources: how cohorts were matched, how transcripts were coded, how schedules were counted, what was excluded and why. It is written so that a reader could repeat the procedures to check the results.

DATA ANALYSIS What the data show, supported by the chapter's figures and tables. This subsection is descriptive; interpretation follows in later subsections.

SUMMARY OF FINDINGS An overarching answer to the evaluation question as stated in operational terms and drawn from the chapter's key findings.

KEY FINDINGS The numbered findings, each with an identifier such as V.4-F1, the source it is based on, the pillar it addresses.

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.1 · EQ 1: English Language Proficiency (ELP)

☀ 4 bright spots in this chapter: V.1-F1 · V.1-F2 · V.1-F3 · V.1-F5

EVALUATION QUESTION	To what extent are students developing English <u>language proficiency</u> , and are <u>English Learners</u> progressing toward <u>reclassification</u> ? (RFP Sec. C)
IN OPERATIONAL TERMS	<i>How much English <u>language proficiency</u> did <u>English Learners</u> gain each year from spring 2023 to spring 2025, measured on the <u>WIDA ACCESS for ELLs</u> composite <u>scale score</u>, and how many students were <u>reclassified out of English Learner</u> status over the same span against the <u>Illinois 4.8 criterion</u>?</i>
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillar: <u>Bilingualism</u> & <u>Biliteracy</u> Strand 4: Assessment & Accountability

DATA SOURCES

QUANTITATIVE	QUALITATIVE
• <u>ACCESS for ELLs</u> (WIDA), spring 2023, 2024 and 2025 • WIDA Research Report RR-2026-3 (Poole & Sahakyan, March 2026) • District <u>Reclassification</u> data	• Administrator interviews • Teacher focus groups • Classroom observations • Parent focus groups

METHODOLOGY

Reporting conventions applied throughout. Growth is reported in composite scale scores and standing in proficiency levels, and the two are never combined. No effect sizes appear in the narrative or in any exhibit. Reclassification counts for spring 2025 are estimates pending confirmation against district reclassification records. Qualitative evidence is aggregated.

DATA ANALYSIS

EQ 1 · ENGLISH LANGUAGE PROFICIENCY · 2.3 DATA ANALYSIS ACCESS for ELLs · spring 2023 · 2024 · 2025

English Language Proficiency: what the ACCESS data show

WHO WAS TESTED

The students behind the numbers

373 distinct students took ACCESS at least once, K–5

All students who tested at least once

373 students

Followed in all three springs — the growth cohort

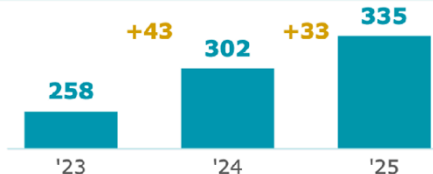
119

Everything reported about growth describes these **119 students**.

COMPOSITE GROWTH

How far students moved

+76 scale-score points gained over two years (n = 119)



Composite scale score, matched cohort

WHEN GROWTH HAPPENED

By grade cohort

Cohort	'23 → '24	'24 → '25	Total
K → 1 → 2	+92	+26	+117
1 → 2 → 3	+42	+21	+63
2 → 3 → 4	+12	+54	+66
3 → 4 → 5	+26	+32	+58

Each row follows the same students as they advance one grade per year. Interval gains n = 118; totals n = 119.

BALANCE ACROSS MODALITIES

Oral and literacy together

+77 / +75 oral · literacy

Oral 264 → 341

Literacy 256 → 332

Neither modality lags the other. English literacy is not developing more slowly than oral language.

MOVING UP THE SCALE

Where students stand

Proficiency level

- 1
- 2
- 3
- 4
- 5
- 6



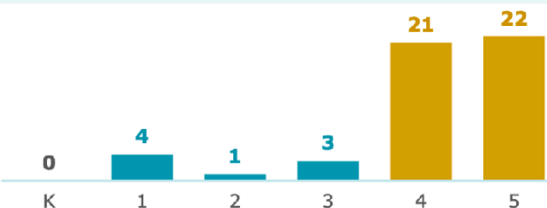
At or above the 4.8 exit level in 2024–25:

19.3% of the 119-student cohort

EXITS FROM EL SERVICES

Reaching the 4.8 level

51 students reached 4.8 and stopped testing · 13.7% of the 373



Grade at exit. 43 of the 51 exits — 84% — occur in grades 4 and 5; none in kindergarten.

Figure V.1.1. English language proficiency: what the ACCESS data show. Matched three-year cohort, n = 119, drawn from 373 distinct testers. Annual growth, SD 112 against the WIDA Consortium.

FINDINGS SUMMARY

119 English Learners took the ACCESS test in all three years and were measured how far each one moved from their own starting point: they gained 76 scale-score points in two years, three of the four grade cohorts matched or beat the national growth benchmark for their grades, and by 2024-25 nearly a fifth had reached the Illinois reclassification level. Because only English Learners take ACCESS, these results describe English Learners' proficiency growth and do not compare dual language with other programs.

KEY FINDINGS

ID	FINDING	DATA SOURCE	DUAL LANGUAGE PILLAR
 V.1-F1	<u>English Learners</u> in the two dual language schools made large, sustained composite growth across three springs.	<u>ACCESS</u> matched <u>cohort</u> , 2023–2025	<u>Bilingualism & Biliteracy</u>
 V.1-F2	Growth was balanced across Oral language and Literacy	<u>ACCESS</u> matched <u>cohort</u> (Figure V.1.1)	<u>Bilingualism & Biliteracy</u>
 V.1-F3	Benchmarked against WIDA Consortium averages, three of the four grade <u>cohorts</u> met or exceeded the matching consortium gain across the two-year window while the oldest <u>cohort's</u> gain was smaller	<u>ACCESS</u> annual pairs benchmarked to WIDA RR-2026-3 (Table V.1a)	<u>Bilingualism & Biliteracy</u>
 V.1-F5	By 2024-25, 19.3 percent of the matched <u>cohort</u> had reached the Illinois 4.8 <u>reclassification</u> criterion,	<u>ACCESS</u> matched <u>cohort</u> (Figure V.1.1)	<u>Bilingualism & Biliteracy</u>
V.1-F7	Students who reach <u>reclassification</u> are primarily in the upper grades of the program. In dual language, these students remain in the program but are no longer classified as ELs.	<u>ACCESS</u> file, 4.8 criterion plus non-testing (Figure V.1.1)	Assessment & Accountability
V.1-F12	Two sessions, one at each site with no shared participants, volunteered that <u>English Learners'</u> English proficiency has improved considerably and that more children are communicating in English. Conversely, they shared a concern that English-dominant students' Spanish has not kept pace.	Teacher focus groups (Table V.1c)	<u>Bilingualism & Biliteracy</u>
V.1-F13	Teachers, parents and students reported that English test data receive the most attention.	Teacher focus groups; parent focus groups (Table V.1c)	Assessment & Accountability
V.1-F14	<u>Reclassification</u> , EL vs. former EL status, was not evident on the classroom rosters.	Dual Language Classroom rosters	Assessment & Accountability

KEY FINDINGS			
ID	FINDING	DATA SOURCE	DUAL LANGUAGE PILLAR
V.1-F15	Classroom observation showed that English is prominent across language blocks. In Spanish-designated blocks students frequently answered in English and teachers accommodated by translating rather than redirecting.	Classroom observations, 38 rubrics (Table V.1c)	<u>Bilingualism</u> & <u>Biliteracy</u>

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.2 · EQ 2: Spanish Language Proficiency (SLP)

☀️ 2 bright spots in this chapter: V.2-F3 · V.2-F8

EVALUATION QUESTION	To what extent are all students developing Spanish <u>language proficiency</u> ? (RFP Sec. C, D)
IN OPERATIONAL TERMS	What do grade 4–5 Spanish writing samples, classroom observations, curriculum review, and teacher, parent, and student focus groups show about students' Spanish proficiency and about the conditions producing it?
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillars: <u>Bilingualism</u> & <u>Biliteracy</u> , <u>Academic Achievement</u> , Strand 4: Assessment & Accountability

DATA SOURCES	
QUANTITATIVE	QUALITATIVE
- Seven grade 4–5 writing samples (four Spanish, three English) scored against the district's grade-4 Spanish <u>Language Arts rubrics</u>. <u>Marco DALE</u> proficiency-level estimates for the same four Spanish samples, across discurso, oración, and palabra/frase. - Curricular materials review: seven entries scored on a seven-dimension <u>rubric</u>, grades 1–4,	- Thirty-eight classroom observations Six student focus groups roughly 30 students. - Six parent interviews: four conducted in English, two in Spanish. - The district's 2026–27 dual language professional learning plan - K-5 Science and Social Studies schedules/ calendars

METHODOLOGY

Seven writing samples (four Spanish, three English) were collected from grade 4–5 classrooms and scored against the district's own grade-4 Spanish Language Arts rubrics (Organization, Development, Language & Conventions), which measure students against the district's stated Spanish Language Arts expectations. The Paridad evaluation team scored four Spanish samples a second time against WIDA's Marco DALE Spanish proficiency-level descriptors, a six-level continuum applied across three dimensions (discurso, oración, palabra/frase). The district rubric asks whether the work meets the district's grade-level Spanish Language Arts standard, whereas the DALE descriptors outline 6 proficiency levels of Spanish-language development.

The figure reports where each Spanish writing sample scored on both instruments. It is a distribution of ratings across two instruments

DATA ANALYSIS				
SAMPLE	GENRE	SLA RUBRIC (ORG / DEV / LANG CONV)	MARCO DALE SLP LEVEL (1-6)	READ TOGETHER
Sample #1	Informational	Meeting / Meeting / Approaching (high)	Level 4 (high)	Near grade <u>standards</u> ; upper-intermediate proficiency
Sample #2	Informational	Meeting / Meeting / Approaching	Level 4 (low)	Meets on structure with sentence-level control
Sample #3	Informational	Meeting / Approaching / Approaching	Level 3-4	Consistent mid-intermediate levels across both s <u>rubrics</u>
Sample #4	Narrative	Meeting / Meeting / Approaching (high)	Level 4 (high)	Near grade <u>standards</u> ; upper-intermediate proficiency

Table V.2.1. DALE proficiency-level estimate and district SLA rubric, by writing sample (n = 4 Spanish samples).

In the materials review, Spanish content quality was rated strong for authentic Spanish, academic register, Latin American authorship, and parity with the English counterpart. However, the review found a time conflict in the schedule: Spanish foundational skills require 40 minutes a day. The district provides a 60 minute block as well as a 60-minute English block on alternating weeks; Spanish science includes enough material for 45–60 minutes a day of instruction; the district has scheduled Spanish social studies and Spanish science 30 minutes a day alternating between science and social studies units weekly which are often lost to foundational skills instruction. The single highest-rated material in the review is the Spanish math supplement, whose 20-minute daily requirement matches the time the schedule protects.

FINDINGS SUMMARY

A small sample of writing samples in English and Spanish from grades 4 and 5, shows students' ability to organize their writing in Spanish. The scoring rubric used placed every writing sample in the intermediate range.

KEY FINDINGS			
ID	FINDING	DATA SOURCE	PILLAR
V.2-F1	Sampled Spanish writing (n = 4) are at an intermediate level, as interpreted by two independent scoring <u>rubrics</u> , the district's SLA <u>rubrics</u> and WIDA's <u>MARCO DALE</u> descriptors, yielding an identical rank order across both. <i>Organization</i> and <i>discurso</i> are consistent strengths; Language & Conventions and oración are at the upper level of performance.	Writing sample analysis of SLA; Marco DALE SLD comparison; teacher focus groups; parent interviews	<u>Bilingualism & Biliteracy</u>
V.2-F2	Seven independent sources agree that daily <u>language allocation</u> runs substantially more English-heavy than in the stated 50%English/50% Spanish design, clustering near 70% English/30% Spanish in instructional practice.	Master-schedule analysis; classroom observations; teacher, student, and parent report; materials review	<u>Academic Achievement</u>
 V.2-F3	Sustained academic Spanish was observed during blocks designated for Spanish instruction (SLA, Esquina de Número). During English instructional time, teachers occasionally used Spanish in the form of isolated vocabulary, immediate translation, or behavior management.	Classroom observations	<u>Bilingualism & Biliteracy</u>
V.2-F4	Consistent Spanish language peer-to-peer and teacher-student interaction (listening and speaking) was observed in a few instances. Observers recorded only a few instances of English-speaking students using Spanish during instruction and they responded in English even when Spanish was the language of the lesson.	Classroom observations; teacher focus groups; parent interviews	<u>Bilingualism & Biliteracy</u>
V.2-F5	The main constraint is instructional time; material quality is generally strong. Spanish-medium materials were rated strong on authenticity and academic register nearly everywhere they were reviewed, while the same reviews record that the minutes each curricular program (in both languages) requires (for instruction and preparation, especially science) exceed the stated minutes in the schedule.	Curricular materials review	<u>Academic Achievement</u>
V.2-F6	Explicit cross-linguistic instruction is limited in both the materials and the observed practice. Observations record that cross-linguistic analysis occurred primarily at the word, specifically the cognate level	Curricular materials review; classroom observations	<u>Bilingualism & Biliteracy</u>
V.2-F7	Accountability is focused on the English measures (<u>language proficiency</u> and <u>academic achievement</u>).	Teacher focus groups; EQ 4 assessment inventory	<u>Bilingualism & Biliteracy</u>

KEY FINDINGS			
ID	FINDING	DATA SOURCE	PILLAR
 V.2-F8	Students hold a positive orientation toward <u>bilingualism</u> , expressed as pride, identity, and career aspiration. Instruction in Spanish is largely undifferentiated between heritage, EL, and English-dominant learners	Student focus groups; parent interviews; teacher focus groups; materials review	<u>Sociocultural Competence</u>

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.3 · EQ 3: Academic Achievement

☀️ 3 bright spots in this chapter: V.3-F1 · V.3-F6 · V.3-F14

EVALUATION QUESTION	What is the achievement of EL students in the Dual Language program for <u>Language Arts</u> and Mathematics? (RFP Sec. C)
IN OPERATIONAL TERMS	What do the data show about achievement in the dual language program on the annual state test, and the interim measures for mathematics and <u>language arts</u> separately, English and Spanish separately, across the three years for which data exist?
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillars: <u>Academic Achievement</u> , <u>Bilingualism</u> & <u>Biliteracy</u> , <u>Sociocultural Competence</u> Strand 4: Assessment & Accountability

DATA SOURCES

QUANTITATIVE	QUALITATIVE
• <u>IAR</u>: ELA and Mathematics, grades 3–5, Spring 2025. • <u>MAP</u>: English reading, mathematics and Reading-Spanish, 2022–23 through 2024–25. • <u>FastBridge</u> earlyReading English (K–1) and CBMreading English, three years — ECRA/LGM Overall Growth Summary, three years • <u>IAR</u> growth dashboards, district level, grades 4–8, All and <u>English Learners</u>.	• Classroom observations • Teacher focus groups • Instructional materials review: Curriculum Materials Evaluation Findings Report • Reading and Writing <u>pacing calendars</u>: six grade-level HMH dual language ELA/SLA calendars • Elementary K-5 Science and Social Studies overview document • Administrator interviews • <u>Language allocation</u> analysis, SY 2025–26.

METHODOLOGY

Quantitative analysis drew on three assessment systems. NWEA MAP results in English language arts, mathematics, and Reading-Spanish were assembled for 2022–23 through 2024–25 and reported two ways: as a three-year series for grades 3 through 5, the only grades tested in all three years (n = 216–247 per subject-year, both sites pooled and weighted by tested students), and as a single-year picture of grades 1 through 5 for 2024–25 (n = 69–89 per cell). One cohort (grade 3 in 2022–23, grade 4 in 2023–24, grade 5 in 2024–25) was followed across all three years at the same two schools (n = 66–89 by subject and year). Illinois Assessment of Readiness results for Spring 2025 were taken from de-identified individual reports for grades 3 through 5 and compared against printed state averages and Illinois cut scores. FastBridge earlyReading and

CBMreading, together with the ECRA/LGM Overall Growth Summary and its Dual Language and Former EL contrasts, supplied the K–1 and district-level material.

Reporting conventions. Every figure separates status from growth and names its comparison group. Growth against national norms is reported as the published percentile and ECRA growth by its published band label, with underlying values carried in the appendix. No effect sizes appear in the narrative or the graphics. Results are pooled across the two sites; no site-level disaggregation is presented.

Qualitative analysis drew on nine substantive teacher focus group transcripts from both sites (April 2026), administrator interviews, classroom observations, parent focus groups, the Curriculum Materials Evaluation Findings Report, the six grade-level HMH dual language ELA/SLA pacing calendars and the K–5 science and social studies overview documents, and the SY 2025–26 language allocation analysis. Transcripts were coded to the themes reported at Figure V.3.5. All qualitative evidence is reported in the aggregate; quotations carry no attribution beyond site, and only where site is relevant.

DATA ANALYSIS

Reporting conventions applied throughout. Each figure separates status from growth and names its comparison group. No effect sizes appear in the narrative or the graphics; ECRA’s growth metric is reported by its published band label. No site-level disaggregation is presented.

SUBJECT	YEAR	TESTED	GROWTH %ILE	VS 50TH NORM	STATUS: FALL	STATUS: SPRING
Mathematics (English)	2022–23	223	55.3	+5.3	69.0	66.6
Mathematics (English)	2023–24	236	52.8	+2.8	64.2	65.1
Mathematics (English)	2024–25	247	62.2	+12.2	61.0	68.0
<u>Language arts</u> (English)	2022–23	222	62.7	+12.7	62.1	71.4
<u>Language arts</u> (English)	2023–24	234	52.0	+2.0	63.6	63.3
<u>Language arts</u> (English)	2024–25	247	60.8	+10.8	60.3	66.3
<u>Language arts</u> (Spanish)	2022–23	216	31.1	–18.9	39.6	31.1
<u>Language arts</u> (Spanish)	2023–24	235	36.0	–14.0	33.8	29.8
<u>Language arts</u> (Spanish)	2024–25	246	36.5	–13.5	29.7	25.8

Table V.3.1. Growth and status on NWEA MAP in English language arts and mathematics in English, three years, dual language program. Comparison group: national norms. Grades 3–5 only: the grades tested in all three years. Both sites pooled and weighted by tested students; n = 216–247 per subject-year.

GRADE	MATHEMATICS (ENGLISH)	LANGUAGE ARTS (ENGLISH)	LANGUAGE ARTS (SPANISH)
Grade 1	growth 71 · status 70→75 (n=72)	growth 66 · status 63→71 (n=72)	growth 39 · status 37→34 (n=69)
Grade 2	growth 52 · status 64→56 (n=75)	growth 65 · status 53→58 (n=75)	growth 31 · status 30→21 (n=74)
Grade 3	growth 58 · status 59→68 (n=79)	growth 71 · status 55→68 (n=79)	growth 26 · status 35→28 (n=79)
Grade 4	growth 61 · status 68→72 (n=79)	growth 56 · status 69→72 (n=79)	growth 42 · status 33→30 (n=79)
Grade 5	growth 66 · status 57→64 (n=89)	growth 56 · status 57→59 (n=89)	growth 41 · status 22→20 (n=88)

Table V.3.2. The same measures by grade, 2024–25, the one year covering grades 1 through 5. Comparison group: national norms. n = 69–89 per cell.

SUBJECT	GRADE 3 · 2022–23	GRADE 4 · 2023–24	GRADE 5 · 2024–25
Mathematics (English)	growth 68 · status 66→68	growth 65 · status 66→71	growth 66 · status 57→64
<u>Language arts</u> (English)	growth 63 · status 54→65	growth 60 · status 68→75	growth 56 · status 57→59
<u>Language arts</u> (Spanish)	growth 42 · status 31→33	growth 38 · status 41→34	growth 41 · status 22→20

Table V.3.3. One cohort followed for three years: the students in grade 3 in 2022–23, grade 4 in 2023–24 and grade 5 in 2024–25. The same grade cohort at the same two schools, n = 66–89 by subject and year.

GRADE	SUBJECT	n	STATUS: MEAN SCALE SCORE	STATE AVG (PRINTED)	STATUS: % PROFICIENT+	GROWTH: MEDIAN SGP
3	ELA	72	750.7	730	68.1%	baseline yr
3	Math	72	751.1	732	63.9%	baseline yr
4	ELA	76	746.8	736	67.1%	65.0
4	Math	76	740.2	732	51.3%	58.0
5	ELA	75	753.3	739	73.3%	48.5
5	Math	75	742.7	728	52.0%	51.5

GRADE	SUBJECT	n	STATUS: MEAN <u>SCALE SCORE</u>	STATE AVG (PRINTED)	STATUS: % PROFICIENT+	GROWTH: MEDIAN <u>SGP</u>
3-5	ELA	223			69.5%	54.0
3-5	Math	223			55.6%	54.0

Table V.3.4. IAR Spring 2025, dual language program. Comparison group: printed state averages and Illinois cut scores. Grades 3-5, both sites pooled

THEME	WHAT TEACHERS REPORT
Delivered allocation	Across both sites and every grade band represented, teachers describe practice as more English-heavy than the 50/50 design which is commonly estimated at about 70 percent English to 30 percent Spanish,
Mathematics	Mathematics is described as predominantly English-medium from the earliest grades, with Spanish appearing mainly in supplementary practice; Spanish-medium mathematics was reported more consistently in the upper elementary transcripts
Assessment	Teachers repeatedly flagged that the high-stakes measures (<u>IAR</u> , <u>MAP</u> Reading and Mathematics in English, and <u>WIDA ACCESS for ELLs</u>) are English-only and inform instructional attention and accountability pressure, while the Spanish measure <u>MAP Lectura</u> was not looked at

Table V.3.5. What teachers report, by theme. Comparison group: none; these are teacher perceptions rather than achievement data. 9 substantive transcripts, both sites, April 2026. Reported in the aggregate; quotations carry no attribution beyond site where site is material.

WHAT WAS EXAMINED	WHAT WAS NOTED FROM THE MATERIALS REVIEW
Quality of the Spanish materials	iArriba La Lectura! (Spanish ELA) scored well on Spanish-language quality and on <u>translanguaging</u> pedagogy, among the strongest materials in the entire review. The materials included Spanish reading passages of high quality and authentically authored; appropriate for use in dual language classrooms. Number Corner, the Spanish mathematics supplement, scored well on quality of the materials and of the inclusion of authentic Spanish language
Coordination between the paired series	Cross-program coordination, how well the paired English and Spanish series align in pacing and build explicit cross-language analyses was not as strong. Thus, the two monolingual programs require coordination between the pairs as it is not explicitly built into the materials.
Time the materials require	Both foundational-skills programs, English and Spanish, were repeatedly flagged as including many more components than fit into the schedule. Following the district's coordinated foundational skills map would reduce the total number of minutes, and allow time for science and social studies, writing across the curriculum and time to read grade level texts in both languages each day.

WHAT WAS EXAMINED	WHAT WAS NOTED FROM THE MATERIALS REVIEW
Cultural and community representation	Except for the HMH Arriba La Lectura, curricular materials require infusion of authentic cultural references and contexts.

Table V.3.6. The review of curricular materials

SOURCE	WHAT IT SHOWS
IAR growth dashboards, district, grades 4–8	Median growth percentile for <u>English Learners</u> rose across three years: 46.0 in 2023, 51.2 in 2024, 53.9 in 2025, against an all-student 45.5, 49.9 and 57.9. <u>English Learners</u> moved from below the national median to above it.
ECRA/LGM, Former EL contrast, 2024–25	Students flagged Former EL showed higher growth than their non-flagged peers in both subjects, and are the only group in the model to reach the <i>Higher than Expected</i> band in ELA-Reading.
WIDA ACCESS for ELLs longitudinal analysis (EQ 1)	Substantial English proficiency growth in the matched <u>cohort</u> , with Oral language and Literacy composites rising in close parallel.

Table V.3.7. The English Learner strand, what the evidence does and does not support. Every row is a district-level or district figure; none isolates English Learners inside the dual language program.

FINDINGS SUMMARY

Dual language students outgrew national peers who began at the same score in English language arts and mathematics in every one of the three years examined, and on the Spring 2025 IAR the program exceeded the printed state average in every tested grade and subject. Growth on MAP Reading in Spanish trailed national peers in all three years, and Spanish Language Arts remains the only Spanish measure the district administers on a standardized basis. These findings were triangulated by administrator and teacher focus groups as well as observations.

Mathematics is not measured in Spanish with a standardized instrument. Math is primarily taught in English and assessed using IAR and NWEA mathematics.

KEY FINDINGS

ID	FINDING	DATA SOURCE	DUAL LANGUAGE PILLAR
 V.3-F1	Dual language students grew faster in English than national peers who began at the same score in English, in all three years and in both English	<u>MAP</u> Growth, three years (Table V.3.1)	<u>Academic Achievement</u>

ID	FINDING	DATA SOURCE	DUAL LANGUAGE PILLAR
	<u>language arts</u> and mathematics in English.		
V.3-F3	Dual language students grew more slowly than national peers who began at the same score, in all three years in Spanish language arts (reading).	MAP Reading-Spanish, three years (Tables V.3.1–V.3.2 and Figure V.3.3)	Bilingualism & Biliteracy
V.3-F4	The district’s only standardized test in Spanish is for Spanish <u>Language Arts</u>	District Data Tracker v7; Curriculum & Pacing Framework; administrator interviews (6 of 6); teacher focus groups	<u>Bilingualism</u> & <u>Biliteracy</u>
 V.3-F6	In the Spring 2025 <u>IAR</u> , the dual language program exceeded the printed state average in every grade and subject.	<u>IAR</u> Spring 2025 individual reports, de-identified (Figure V.3.4)	<u>Academic Achievement</u>
V.3-F8	Teachers describe Spanish reading and Spanish <u>academic achievement</u> as falling short while English outcomes continue to grow. During observations in Spanish-medium instruction, students answered in their dominant language rather than the target language.	Teacher focus groups ; classroom observations ; <u>language allocation</u> analysis	<u>Bilingualism</u> & <u>Biliteracy</u>
V.3-F9	Mathematics is taught in both languages but assessed only in English. The daily program pairs a 60-minute Math block in English with a 30-minute Number Corner block in Spanish at every grade, while both <u>IAR</u> and <u>NWEA</u> mathematics are administered in English. Two-thirds of designed mathematics time is therefore in the language of the test and one-third is not assessed on a standardized measure in Spanish.	<u>Language allocation</u> analysis SY 2025–26; Curriculum & Pacing Framework; classroom observations	<u>Academic Achievement</u>
V.3-F13	Different stakeholder groups described the program’s emphasis on English measures and English outcomes for ELs. Spanish <u>language development</u> is	Teacher focus groups (Figure V.3.5); administrator interviews;	Assessment & Accountability

ID	FINDING	DATA SOURCE	DUAL LANGUAGE PILLAR
	not assessed with a standardized measure.	District Data Tracker v7; parent focus groups	
 V.3-F14	The Spanish curriculum materials are of high quality across all content areas.	Curriculum Materials Evaluation parent focus groups	Bilingualism & Biliteracy

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.4 · EQ 4: Biliteracy Assessment Framework

EVALUATION QUESTION	How might a comprehensive <u>biliteracy</u> assessment framework be conceptualized for the dual language program in SD 112? (RFP Sec. C, D)
IN OPERATIONAL TERMS	What data and what instrumentation are in place to capture dual language students' literacy and oral language (listening and speaking) development in English and in Spanish, and across both languages?
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillar: <u>Bilingualism</u> & <u>Biliteracy</u> Strand 4: Assessment & Accountability

DATA SOURCES

QUANTITATIVE	QUALITATIVE
- K–5 curriculum assessment inventory: District assessment verification worksheet: Spanish and English writing samples, grades 4–5 - Classroom observation <u>rubric</u>	- Teacher focus groups: - Parent focus groups and interviews - Instructional materials review

METHODOLOGY

This question asks what the district has in place to assess student literacy in two languages. The method is an inventory of measures with gap analysis. The inventory drew on two sources: 1) the district and state measures named in the EQ 4 matrix, and 2) The assessment components published for the eight adopted K–5 programs — Into Reading, ¡Arriba la Lectura!, CKLA Skills, Caminos Lectoescritura, Bridges in Mathematics, Number Corner, Social Studies Alive!, and FOSS Science producing 91 items.

Only the teacher protocol and the materials review asked about assessment directly; parents and students raised it unprompted inside answers to other questions. All qualitative evidence is reported in the aggregate.

DATA ANALYSIS

SECTION V · V.4 · EQ 4 BILITERACY ASSESSMENT FRAMEWORK

Biliteracy Assessment Inventory, K–5: The Measures in Use

Only measures confirmed in use are shown; dashed boxes mark what is not measured. Complete district verification, August 2026.

☐ In use districtwide ☐ In use at teachers' discretion ☐ Not in place

n = 91 documented assessment components (81 embedded in the 8 adopted programs; 10 district, state, and local measures), plus FastBridge, which appears in no publisher documentation and is given in both languages. The district answered 89 of 91: 28 in use, 17 partial or at teacher discretion, 44 not in use. Unanswered: CKLA Skills End-of-Year Assessment; FOSS I-Checks. The 44 not in use are omitted here by design and reported in Table V.4a.

Area of assessment	English	Spanish	Both / cross-linguistic
Language proficiency (oral + written)	WIDA ACCESS for ELLs <i>annual · listening, speaking, reading, writing · English learners only</i> WIDA Screener <i>at entry · new enrollees, K–5</i> Interim English language proficiency measure	Spanish language proficiency measure	Relationship between the two languages
Reading & literacy	NWEA MAP — Reading <i>3×/yr · interim, norm-referenced growth</i> IAR — ELA (Gr 3–5) <i>annual · summative (state)</i> Into Reading — weekly & module <i>3 of 15 components in use; 1 partial</i> CKLA Skills (K–2 only) <i>placement, unit, fluency — 4 of 14 in use; 2 partial</i>	NWEA MAP Lectura (K–5) <i>2–3×/yr · interim · not normed on English-speaking students</i> iArriba la Lectura! — weekly & module <i>2 of 9 components in use; 1 partial</i> Caminos Lectoescritura (K–2 only) <i>unit assessments — 1 of 6 in use; 2 partial</i>	—
Screening & early literacy (K–2)	FastBridge <i>universal screener, diagnostic & progress monitoring</i>	FastBridge Español <i>early literacy screening in Spanish, parallel to English</i>	—
Writing	Into Reading writing + constructed-response rubrics <i>in use · per module</i> Local English writing sample (Gr 4–5) <i>4 pieces/yr · no anchor set or double-scoring reported</i>	Local Spanish writing sample (Gr 4–5) <i>4 pieces/yr · iArriba la Lectura! Pautas de calificación not in use; no anchor set or double-scoring reported</i>	—
Oral language / oracy	WIDA ACCESS — Listening & Speaking <i>annual · English learners only</i>	Spanish oral language proficiency measure	—
Mathematics	NWEA MAP — Mathematics <i>interim, norm-referenced growth</i> IAR — Mathematics (Gr 3–5) <i>annual · summative (state)</i> Bridges in Mathematics, 2nd Ed. <i>9 of 10 components in use</i>	Mathematics measure in Spanish	—
Content — Science & Social Studies	—	Social Studies Alive! <i>no component in districtwide use; 8 of 11 at teacher discretion</i> FOSS Science <i>no component in districtwide use; 3 of 10 at teacher discretion</i> Content assessment of record	—
Sociocultural competence	—	—	Measure in either language

Paridad Education Consulting · SD 112 Dual Language Program Evaluation · Figure version September 28, 2026

Source: district verification of the K–5 assessment inventory, August 2026

Figure V.4.1. The measures in use, K–5, by area, language, and purpose, following complete district verification.

FINDINGS SUMMARY

The district measures reading and foundational literacy in both languages from kindergarten through grade 2, and English proficiency in all four domains from K–5, but has no Spanish language proficiency measure. Teachers, parents and students report that English results are acted on more readily than Spanish results.

KEY FINDINGS			
ID	FINDING	DATA SOURCE	PILLAR
V.4-F1	English <u>language proficiency</u> is measured annually across all four domains by <u>WIDA ACCESS</u> , with the <u>WIDA Screener</u> at entry. No Spanish <u>language proficiency</u> measure is in use.	Verification worksheet; <u>WIDA ACCESS</u> and <u>WIDA Screener</u>	<u>Bilingualism & Biliteracy</u>
V.4-F2	Spanish <u>language arts</u> reading and foundational literacy are broadly comparable to English from kindergarten-5. <u>NWEA MAP</u> Lectura has not been normed on English speaking students.	District verification;	<u>Bilingualism & Biliteracy</u>
V.4-F3	Spanish and English writing are both collected four times a year at grades 4–5.	District verification; document review	<u>Bilingualism & Biliteracy</u>
V.4-F4	The relationship between the two languages is not captured; such as: cross-linguistic or cross-cultural transfer, <u>translanguaging</u> , or multimodal expression.	Verified inventory	<u>Sociocultural Competence</u> · <u>Bilingualism & Biliteracy</u>
V.4-F5	<u>Academic achievement</u> in Spanish is measured in reading on an interim basis. No annual standardized Spanish <u>language arts</u> measure; <u>NWEA</u> Mathematics in Spanish exists, however, mathematics is generally taught in English.	District verification; <u>NWEA MAP</u> Lectura; curriculum inventory; materials review	<u>Academic Achievement</u>
V.4-F6	Annual Spanish <u>language proficiency</u> test and interim English <u>language proficiency</u> assessment are not in place.	Instrument market review; district verification	<u>Bilingualism & Biliteracy</u>
V.4-F7	In general, standardized Spanish assessment results are not used in instructional decisions, while English results are.	Teacher and parent focus groups (16 sessions)	<u>Academic Achievement</u> · <u>Bilingualism & Biliteracy</u>
V.4-F8	Assessment volume is high. Teachers describe a dense calendar (<u>interim assessments</u> , module tests, <u>MAP</u> , <u>IAR</u> , and common assessment) with little discussion of results. Students report Spanish blocks moved or dropped when testing in English is scheduled.	Teacher focus groups; student focus groups; instructional materials review	<u>Bilingualism & Biliteracy</u> · <u>Academic Achievement</u>
V.4-F9	In the district's instructional materials review, no adopted material rated higher than Partial on culturally & linguistically inclusive assessment.	Instructional materials review (8 usable responses of 10)	<u>Bilingualism & Biliteracy</u>

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.5 · EQ 5: Class Size and Composition

 2 bright spots in this chapter: V.5-F1 · V.5-F9

EVALUATION QUESTION	What are the optimal class sizes and student composition ratios for K–5 dual language classrooms in SD 112? (RFP Sec. B)
IN OPERATIONAL TERMS	Do the class sizes and composition ratios currently in place support the two-way model, and what configurations does the evidence indicate would strengthen it?
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillars: <u>Bilingualism</u> and <u>Biliteracy</u> & <u>Sociocultural Competence</u> Strand 1 (Program Structure, supporting Strand 7) Support and Resources

DATA SOURCES

QUANTITATIVE	QUALITATIVE
- Class size and composition roster - Field data collection schedule - Regulatory frame: 23 Ill. Adm. Code Part 228 and 105 ILCS 5/14C, retrieved and verified at source. - Research frame: the class-size and two-way composition literature, retrieved and verified at source. <u>ACCESS longitudinal</u> matched-cohort analysis	- Administrator interviews - Teacher focus groups - Classroom observations - Parent focus groups and interviews - Student focus groups

METHODOLOGY

Two primary procedures produced the evidence in this chapter.

Roster reconciliation. Every dual language section in the district was transcribed from the 2025 by-teacher distribution and reconciled to that document’s printed totals. All figures tie out: 293 students at Oak Terrace, 199 at Red Oak, 492 combined, with IEP and 504 counts matching the source panels.

Frame assembly. Illinois statute and administrative rule, and the empirical literature on class size and two-way composition were retrieved and verified at source.

DATA ANALYSIS

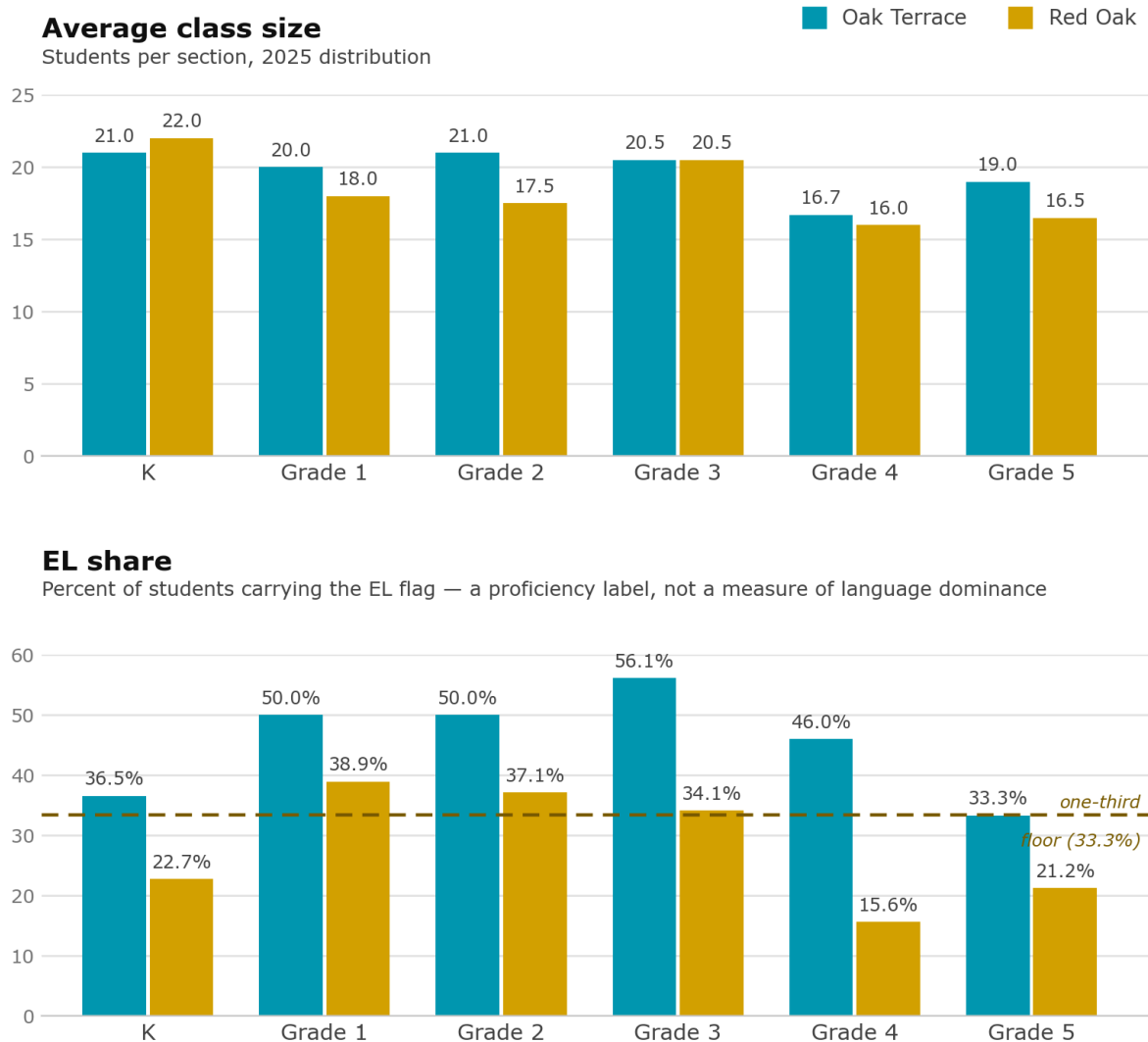
Reporting conventions applied throughout. Qualitative evidence is reported in the aggregate. Site-level detail is presented here in light of the impending merger of the two sites. The site contrast is the finding but Finding 2 guides how any such contrast may be read.

	OAK TERRACE				RED OAK				MERGED	
Grade	Sec.	Students	Average	EL	Sec.	Students	Average	EL	Students	EL
K	3	63	21.0	36.5%	1	22	22.0	22.7%	85	32.9%
1	2	40	20.0	50.0%	2	36	18.0	38.9%	76	44.7%
2	2	42	21.0	50.0%	2	35	17.5	37.1%	77	44.2%
3	2	41	20.5	56.1%	2	41	20.5	34.1%	82	45.1%
4	3	50	16.7	46.0%	2	32	16.0	15.6%	82	34.1%
5	3	57	19.0	33.3%	2	33	16.5	21.2%	90	28.9%
All	15	293	19.5	44.0%	11	199	18.1	29.1%	492	38.0%

Table V.5.1. The program enumerated, 2025 distribution. The table includes all 26 dual language sections and all 492 students (no sampling). Averages are students divided by sections. EL/LEP share is the proportion of that cell identified as English language learners. . Read limitation L1 and V.5-F16 before treating it as composition, because reclassification removes students as they succeed and from the EL category, the upper-grade cells are not measured on the same base as the lower ones. The merged columns are the arithmetic sum of the two sites and reconcile to the program totals of 492 students and 187 EL/LEP (38.0 percent)

Reading the table. Section sizes across the program run from 16 to 23. The two sites are largest and smallest at opposite points of the grade span: Oak Terrace runs at 20 to 21 from kindergarten through grade 3 and drops to 16.7 in grade 4, while Red Oak is largest in kindergarten and grade 3 and runs at 16.0 to 16.5 in grades 4 and 5. Red Oak kindergarten is the only single-section grade anywhere in the program.

A caution on reading down the EL columns. The decline from grade 3 to grade 5 at both sites is not a change in who is enrolled. Reclassification out of English learner status at the Illinois ACCESS threshold drops students from these EL counts permanently. These students keep the same placement and services but are no longer counted as EL. The effect accumulates with each year of enrollment, so the upper-grade cells are systematically lower than the lower-grade cells for reasons that have nothing to do with the language children speak. As a result, the six grade cells in each column are not measured on a common footing and should not be compared with one another as though they were.

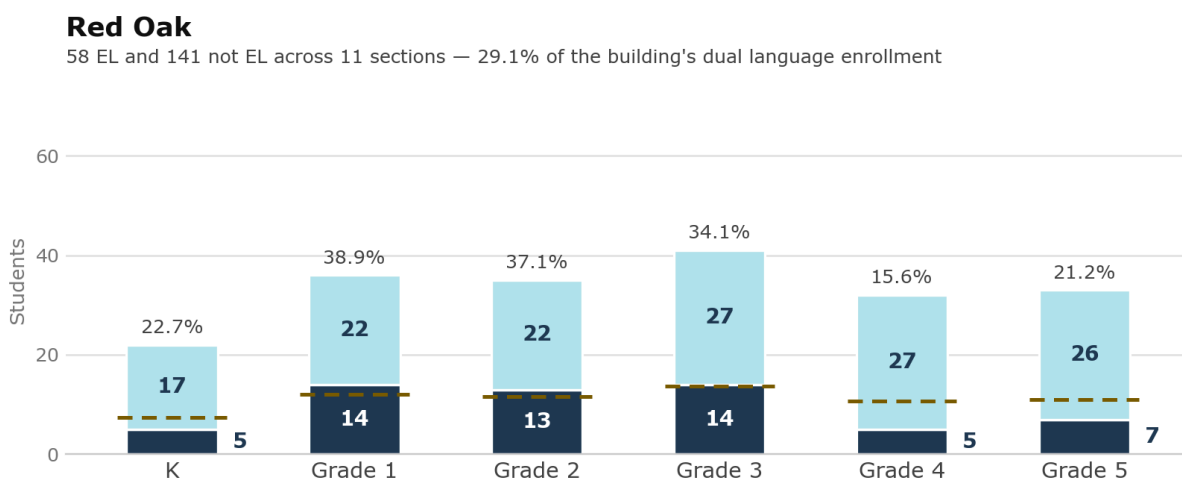
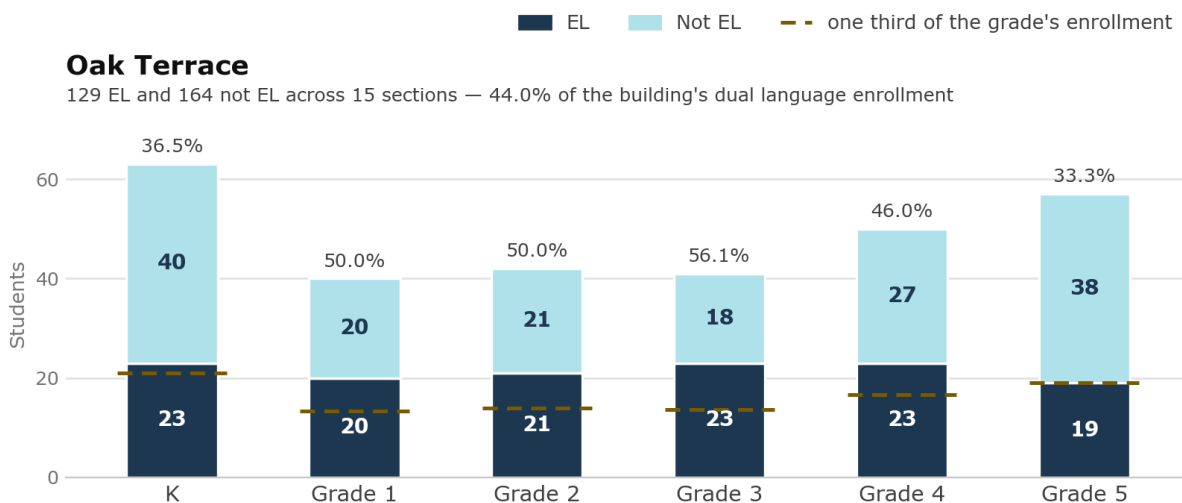


n = 492 students in 26 sections. Source: SD112_ClassSize_Composition_2025.xlsx;
grade assignments resolved in SD112_FieldWeek_DataAudit.docx, Table 2.

Figure version September 28, 2026

Figure V.5.2. Average class size and EL/LEP share by grade and site, 2025 distribution. All 26 sections and 492 students. The dashed line in the lower panel is the one-third floor.

Reading the figure. Class size is roughly flat to declining across the grades while EL share rises through grade 3 and then falls away. The drop off in ELs in the upper grades is explained by reclassification not reflected in the counts used for this graphic (see Finding 8).



Percentages above each bar are the EL share of that grade. Counts tie to site totals (Oak Terrace 293, Red Oak 199, combined 492). EL is a proficiency classification, not a measure of language dominance — a reclassified student remains a native Spanish speaker. See limitation L1. Source: SD112_ClassSize_Composition_2025.xlsx; grade assignments resolved in SD112_FieldWeek_DataAudit.docx, Table 2.

Figure version September 28, 2026

Figure V.5.3. EL and not-EL students by grade and site, 2025 distribution, reported as counts rather than shares. The dashed marker is one third of that grade's enrollment, computed per grade so the floor can be read directly against each bar. Site is distinguished by panel rather than by color; the colors here carry EL status only and do not correspond to the site colors in Figure V.5.2. Counts were reconstructed per grade and reconciled to the site totals of 293 and 199.

FINDINGS SUMMARY

The program's twenty-six sections average 18.9 students. No class exceeds either the regulatory or research upper thresholds. Classes that included high numbers of students with IEPs did not exceed the 70%/30% special rule because these classes included students with pull out speech services only.

KEY FINDINGS

Findings describe what the analysis shows. Interpretation belongs under Interpretation and limitations. Each finding carries an identifier for citation in Sections VI and VII.

ID	FINDING	DATA SOURCE	PILLAR
 V.5-F1	The 25-26 SY dual language student counts by classroom are as follows: 26 sections, 492 students, averaging 18.9 per class. Oak Terrace runs fifteen sections enrolling 293 at an average of 19.5; Red Oak runs eleven enrolling 199 at an average of 18.1. Sections range from 16 to 23.	Class size roster 2025; field schedule (Table V.5.1)	Program Structure
V.5-F2	Class size runs the largest in K-3 where <u>English learner</u> and special education density peaks.	Class size roster 2025 (Figure V.5.2; Table V.5.1)	Program Structure · <u>Academic Achievement</u>
V.5-F3	Red Oak kindergarten is the program's only single-section grade, and the field's design guidance calls for at least two classes per grade.	Class size roster 2025; Howard & Christian (2002)	Program Structure
V.5-F4	Students who are <u>reclassified</u> out of EL status stay in the dual language classroom, so the count on paper falls while the actual makeup of the room is unchanged. The lower numbers of ELs in K and in the upper grades (4 & 5) at the Red Oak site mean two different things. The kindergarten number is reliable because no kindergartner has been <u>reclassified</u> yet, so the EL count is close to the real number of Spanish-speaking children, and it lines up with what the applicant pool shows. The grade 4 and grade 5 numbers are not as accurate.	Class size roster 2025 (Figures V.5.2, V.5.3)	<u>Bilingualism</u> and <u>Biliteracy</u>
V.5-F5	The district continues to augment and improve its language screening at intake to more specifically account for the variety of language learners. Those categories include simultaneous bilingual learners (those students who have been exposed to two languages before the age of 3), or <u>Heritage speakers</u> (those students with Spanish Language and culture backgrounds who may be English proficient), so the composition half of EQ 5 is more complex than the question suggests. EQ 5 asks for composition models, which in a two-way program means the ratio of	District records; teacher focus groups; administrator interviews	<u>Bilingualism</u> and <u>Biliteracy</u>

ID	FINDING	DATA SOURCE	PILLAR
	partner-language speakers to English speakers.		
V.5-F6	A binding regulatory ceiling on dual language class size exists. Illinois rule requires that the student-teacher ratio in the ESL and home-language components of a bilingual program not exceed 90 percent of the building's average general education ratio for the same grades, measured as of September 30 (23 Ill. Adm. Code 228.30(b)(1), amended effective July 1, 2025). Where a building's general education average is 22 in a grade, the ceiling for its bilingual sections in that grade is 19.8. This is the only enforceable class-size number the evaluation has identified: Illinois sets no general elementary cap. One question remains open. The rule never uses the words dual language, two-way or immersion, and does not address how the ratio is computed for a classroom roughly half of whose students are not <u>English learners</u>.	23 Ill. Adm. Code 228.30(b)(1); 105 ILCS 5/14C-6; NSEA agreement Art. 4.3	Program Structure
V.5-F7	The 2025 roster is not the correct planning basis: the program the recommendation covers will be considerably smaller The dataset describes 492 students in 2025. Published district figures record dual language enrollment falling from 695 in 2020–21 to 523 in 2024–25, with projections of 489 for 2025–26 and 447 for 2026–27. A recommendation governing 2027–28 built on the 2025 roster would overstate the program by roughly ten percent before the opening year begins.	Class size roster 2025; published district enrollment figures	Program Structure
V.5-F8	In a dual language program, <u>reclassification</u> changes the label given to the student. They continue to participate in the DL program so the composition variable falls while the classroom does not change at all. Illinois reclassifies a student out of <u>English learner</u> status at an <u>ACCESS</u> composite of 4.8. In an ESL or transitional bilingual program that is an exit: the student leaves the program and the services change. In a dual language program it is not. The student keeps the same seat in the same section, receives the same instruction in the same two languages,	<u>ACCESS</u> longitudinal matched <u>cohort</u> (EQ 1); 23 Ill. Adm. Code 228.25; class size roster 2025	<u>Bilingualism</u> and <u>Biliteracy</u> · Program Structure

ID	FINDING	DATA SOURCE	PILLAR
	and has the same Spanish skills as before. Only the count changes. As a result, the program's single composition variable declines across the grade span for purely administrative reasons and declines fastest in the grades that have had the most years in which to cross the threshold.		
 V.5-F9	The only class-size number any stakeholder group has offered is a ceiling, and the program already runs well below it. Parents ask the district to commit either to a co-teaching model or to a firm cap of 25 to 26 students. Set beside the other figures in this chapter, parents are asking for a guaranteed limit that current class sizes already meet. The roster average is 18.9, no section exceeds 23, observed attendance runs 14 to 19, and the narrative observation overview independently reports 13 to 18. All of these figures are below the ceiling parents cited.	Parent focus groups, 6 Protocol 3 sessions (Table V.5c)	Program Structure · Support and Resources

SECTION V · DATA ANALYSES, REPORTING & FINDINGS

V.6 · EQ 6: Curriculum Coherence and Pacing

☀ 3 bright spots in this chapter: V.6-F3 · V.6-F4 · V.6-F5

EVALUATION QUESTION	How coherent, vertically aligned, and consistently paced is the dual language curriculum across grade levels in SD 112?
IN OPERATIONAL TERMS	Do the scope and sequence, the <u>pacing calendars</u>, the materials, and the <u>language allocation</u> add up to one coordinated K–5 program, or to two programs taught side by side?
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillars: <u>Bilingualism</u> & <u>Biliteracy</u> , <u>Academic Achievement</u> , <u>Sociocultural Competence</u> Strand 2: Curriculum

DATA SOURCES	
QUANTITATIVE	QUALITATIVE
• <u>Language Allocation</u> Analysis • Six master schedules • K–5 Dual Language Coordination Analysis. • Six HMH Dual SLA/ELA <u>pacing calendars</u> for SY 2025–26 • 50/50 Model Time Allocation schedule • Instructional Materials Review.	• Administrator interviews • Teacher focus groups • Student focus groups • Parent interviews • Classroom observations • K-5 Science and Social Studies guides/ calendars

METHODOLOGY

Four procedures produced the evidence in this chapter.

Calendar reconciliation. The six grade-level HMH dual language ELA/SLA pacing calendars for SY 2025–26 were transcribed module by module and the start month of each of the nine core literacy modules recorded by grade, producing Table V.6.1. Coordination was judged on two tests: whether the grades 1-5 units open and close the same module within the same window (kindergarten has a different unit length), and whether the pairing of Spanish-lead or English-lead reading week modules was consistent at every grade.

Allocation analysis. Six master schedules and the 50/50 Model Time Allocation schedule were coded block by block for language of instruction, content area, and minutes, and the delivered allocation was compared against the model the district describes in its Request for Proposal.

Materials review. Eight instructional materials were rated using a tool that was adapted from the ELSF Benchmarks of Quality for Materials on a three-point scale (1 = limited or absent, 2 = partial, 3 = strong) for overall curricular quality and, separately, for three dual language components: cross-linguistic connections, support for biliteracy development, and cultural relevance, producing Tables V.6.2 and V.6.3. Overall quality and dual-language readiness are reported as separate ratings and are not combined into a single score.

Qualitative corroboration. Administrator interviews, teacher and student focus groups, parent interviews, classroom observations, and the K–5 science and social studies guides and calendars were coded for coherence, vertical alignment, and pacing, then read against the calendar and allocation findings. All qualitative evidence is reported in the aggregate.

DATA ANALYSIS

GRADE	M1	M2	M3	M4	M5	M6	M7	M8	M9
Kindergarten	Aug	Oct	Nov	Jan	Feb	Mar	Apr	—	—
Grade 1	Aug	Sep	Oct	Nov	Jan	Jan	Feb	Mar	Apr
Grade 2	Aug	Sep	Oct	Nov	Jan	Jan	Feb	Mar	Apr
Grade 3	Aug	Sep	Oct	Nov	Jan	Jan	Feb	Mar	May
Grade 4	Aug	Sep	Oct	Nov	Jan	Jan	Feb	Mar	May
Grade 5	Aug	Sep	Oct	Nov	Jan	Jan	Feb	Apr	May

Table V.6.1. Core Literacy Module start month by grade, SY 2025–26. Six grades representing the nine-module HMH sequence. Kindergarten is not aligned due to the design of the program at that grade level. Grades 1–5 are designed with longer units.

MATERIAL SUBJECT LANGUAGE GRADE	CROSS LANGUAGE CONNECTIONS	<u>BILITERACY</u>	OVERALL QUALITY	CROSS-CULTURAL CONNECTIONS
Into Reading (HMH) ELA · English · Grade 3	3: Strong	3: Strong	3: Strong	2: Partial
¡Arriba la Lectura! (HMH) ELA · Spanish · Grade 4	3: Strong	2: Partial	3: Strong	2: Partial
CKLA Skills (Amplify)	1: Limited	1: Limited	2: Partial	1: Limited/ Absent

Foundational · English · Grade 1				
Camino Lectoescritura (Amplify) Foundational · Spanish · Grade 1	2: Partial	2: Partial	2: Adequate with support	2: Partial

Table V.6.2. Literacy Materials Evaluation: Overall and dual language-specific component ratings by material (1 = Limited or absent, 2 = Partial, 3 = Strong).

MATERIAL SUBJECT LANGUAGE GRADE	CROSS LANGUAGE CONNECTIONS	<u>BILITERACY</u>	OVERALL QUALITY	CROSS-CULTURAL CONNECTIONS
Bridges in Mathematics Math · English · Grade 2	1: Limited	N/A	3: Strong	1: Limited
Number Corner Math Spanish • Grade 2	2: Partial	3: Strong	3: Strong	1: Limited
Social Studies Alive! (TCI) Social Studies Alive! · Spanish · Grade 3	2: Partial	1: Limited	3: Strong	1: Limited
FOSS Science (Delta) Science · Spanish · Grade 2	2: Partial	2: Partial	3: Strong	2: Partial

Table V.6.3. Mathematics Materials Evaluation: Overall and dual language-specific component ratings by material (1 = Limited or absent, 2 = Partial, 3 = Strong).

Overall quality is reviewed separately from the material’s dual-language readiness: five of eight reviewed materials scored strong as curriculum. The majority of the materials scored partial to strong on the dual language elements and most could be used as designed and a few with some adaptations. Of the seven areas reviewed, cross-cultural and translanguaging pedagogy scored lowest. The curricular materials are all rich in activities, teacher support and instructions. The Dual Language Guide in the HMH materials, the authentic science materials and experiment set up built into the FOSS curriculum, and the supporting manipulatives and explorations in the math curricula are some examples of this excellent instructional support for teachers.

In addition to supporting teachers, the curricular materials also provide quality rigorous learning opportunities for students. This abundance is a strength but also a challenge for the program. Each set of materials requires many more minutes of instruction to complete all the content contained in each unit or module. The clearest example is the K–2 foundational skills pair: CKLA requires 60 minutes per lesson and Caminos is nominally 40 but in practice it often takes 60 minutes, so the pair cannot both be delivered each day. 20-40 minutes of foundational skills a day during the core reading and writing block in the language determined by the pacing guide would allow time for the other content areas. FOSS Science needs approximately 45 to 60 minutes of instruction in addition to the time teachers require to set up and take down an experiment, but often this science is cut short or cut out. The schedule allows for 30 minutes a day for either science or Social Studies. A variety of stakeholders describe social studies and science as the first to be left off the schedule when time runs out.

Number Corner (30 minutes, Spanish) and Math (60 minutes, English) are two fixed-language blocks of a single mathematics program, not a paired sequence moving between languages. Bridges in Mathematics has no Spanish companion in current use; the Spanish supplement currently in circulation was not designed as a companion to it. Both IAR and NWEA mathematics are administered in English.

FINDINGS SUMMARY

The literacy calendar is coordinated in grades 1 through 5, and the schedule pairing odd and even modules with Spanish or English-lead reading week is identical at every grade. The dual language curricula includes separate explicit connections between the two languages within language arts (Dual Language Implementation Guide) No explicit connections are made between number corner (Spanish) and the core mathematics program. Both science and social studies (Spanish) instructional allocation minutes are not fully implemented as published in the curriculum pacing schedule.

KEY FINDINGS

ID	FINDING	DATA SOURCE	PILLAR
V.6-F1	Delivered <u>language allocation</u> leans English-heavy against the stated 50% English/50% Spanish <u>program model</u> . <u>Language allocation</u> appears closer to 59%English/41% Spanish by instructional time and 65% English/35% Spanish across the entire school day; 60% English/40% Spanish by the literacy <u>pacing calendars</u> ' weekly minutes, and roughly 70%English/30% Spanish as teachers, parents, and students independently describe daily practice.	<u>Language Allocation</u> Analysis v2; K–5 Coordination Analysis; teacher, parent, and student focus groups	<u>Bilingualism & Biliteracy</u>
V.6-F2	5 of the 8 reviewed materials received the highest rating on overall quality. Cross-cultural	<u>Language Allocation</u> Analysis v2; classroom	<u>Bilingualism & Biliteracy</u>

	connections and <u>translanguaging</u> pedagogy score lowest of seven areas reviewed on the materials evaluation. The curriculum materials each require more instructional time than the schedule allocates.	observations; student focus groups	
 V.6-F3	Vertical coordination of the literacy calendar is solid. Grades 1 through 5 open and close all nine HMH modules within days of each other across the year, and the odd/even module-to-language-lead rule is applied identically at every grade band.	K-5 Dual Language Coordination Analysis	<u>Academic Achievement</u>
 V.6-F4	The pacing plan has been designed to deliver twomonolingual curricula across content areas. The district has already done some of the work to make connections between curricula: a crosswalk document that connects core literacy and social studies; <u>pacing calendar</u> for core literacy instruction & writing in English and Spanish; <u>pacing calendars</u> for Spanish Science and Social Studies; as well as a map of high priority foundational skills to teach in each language. Professional learning sessions have been delivered related to teaching the K-3 HMH in the most efficient manner.	Teacher focus groups administrator interviews; Instructional Materials Review	<u>Academic Achievement</u>
 V.6-F5	The writing curriculum for grades 4 and 5 has been coordinated in English and Spanish, addresses a range of genres of writing and connects to HMH modules.	Instructional Materials Review	<u>Academic Achievement</u>
V.6-F6	Science and social studies in practice are the first areas of instruction cut under time pressure, and they represent an important block in the program's Spanish side of instruction. Other sources describe the same displacement: a weekly inquiry block shrinking from two or three sessions to about one; content areas of science and social studies as the first thing pulled reducing the amount of science and social studies students receive until grade 3	Teacher focus groups; administrator interviews; Instructional Materials Review; <u>Language Allocation</u> Analysis	<u>Academic Achievement</u>
V.6-F7	Cross-linguistic instruction primarily occurs at the word level in both the materials and the reported practice. Every curricular material receiving a substantive cross-language rating connects at word or cognate level. Parents and students mention the lack of crosslinguistic	Instructional Materials Review; administrator interviews; parent interviews; student focus groups	<u>Sociocultural Competence</u>

	connections made during instruction, and students asked for more of it.		
V.6-F8	Number Corner in Spanish and the Bridges Math in English are two fixed-language blocks of one mathematics program rather than a complementary sequence that moves between languages. Bridges in Mathematics has no Spanish companion in current use, and the Spanish supplement in circulation is not a partner of the program. Both <u>IAR</u> and <u>NWEA</u> mathematics are administered in English.	<u>Language Allocation</u> Analysis; Instructional Materials Review	<u>Academic Achievement</u>
V.6-F9	The program has no defined K–5 progression for oral <u>language development</u> . Teachers report that there are no oracy <u>standards</u> for the dual language program; no statement of what students should be able to produce orally, in each language, at each grade.	Materials Review; Teacher focus groups; classroom observations	<u>Bilingualism & Biliteracy</u>

VI. DELIVERABLES AND RECOMMENDATIONS

Every chapter in this section follows a similar structure. The deliverable and the scope of what was asked in the RFP is stated. The related Evaluation Questions (EQs) are listed with the pertinent findings from the data analysis. The premises the recommendations are based on are stated. Each deliverable may contain an anchor figure and/or proposes a model or framework the district can choose to follow. Each deliverable then lists recommendations proposed by the evaluation team along with some supporting recommendations and guidance for implementation. Each deliverable ends with some cautions on use of the findings and recommendations.

SECTION VI · RECOMMENDATIONS BY DELIVERABLE (RFP SEC. D)

VI.1 · Class Size Recommendations

Deliverable 1

DELIVERABLE	VI.1 · Class Size Recommendations (2027–28).
THE DECISION	What class sizes and speaker-composition ratios are optimal for the Northshore Dual Language program?
SCOPE	Recommendations for class size and speaker composition K–5 SD 112 Dual Language Program
ANCHORED BY	EQ 5 · Class Size and Composition
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillar: <u>Sociocultural Competence</u> , equal status for both languages by design. Strand 1: Program Structure
RFP REFERENCE	RFP Sec. D.

FINDING ID	WHAT IT ESTABLISHES
V.5-F1	The program enumerated: 26 sections, 492 students, averaging 18.9.
V.5-F2	Class size runs the largest in K–3 where <u>English learner</u> and special education density peaks.
V.5-F3	Red Oak kindergarten is the program’s only single-section grade, and the field’s design guidance calls for at least two classes per grade.

FINDING ID	WHAT IT ESTABLISHES
V.5-F4	The shift to lower numbers of EL Spanish speakers from Kindergarten to grade 5 is a result of ELs students who are <u>reclassified</u> out of EL status stay in the dual language classroom, so the count on paper falls while the actual makeup of the room is unchanged.
V.5-F5	The district continues to augment its language screening at intake to more specifically account for simultaneous bilingual learners and <u>heritage speakers</u> , so the composition half of EQ 5 is more nuanced than the question suggests.
V.5-F6	Illinois rule caps the bilingual student-teacher ratio at 90 percent of the building general education ratio for the same grades.
V.5-F7	The 2025 roster is not the correct planning basis as it doesn't account for projections that have student enrollment fall to 447 by 2026–27 school year.
V.5-F8	<u>Reclassification</u> at the <u>ACCESS</u> 4.8 composite changes only a student's label, so the EL count falls as more students reach proficiency.
V.5-F9	Parents ask for a firm 25 to 26 cap or a co-teaching model. The program already operates well below that ceiling.
V.5-F10	Proficiency-balanced grouping is already practiced and appears to be teacher design rather than district practice.

PREMISES

- **Class size and composition need to be decided together.** A ratio target is impacted by a section size, and a section size is influenced by a composition rule. Every recommendation below therefore names both, or names which one it is being referred to and why.
- **The recommended class size is based on instructional needs.** There is no class-size research in dual language or immersion settings. The research does establish an instructional requirement that small homogeneous groups of three to six students, interact daily, for at least thirty minutes. A section size is defensible only as the answer to what makes that schedulable, bounded by the two enforceable ceilings in the Illinois rule (bilingual ratio in Part 228 and the limit in Part 226) on the share of a general education class carrying an IEP. No research supports presenting any figure as an optimum.

GRADE	OAK TERRACE	RED OAK	COMBINED	SECTIONS TODAY	AT 24 UPPER LIMIT
K	63	22	85	4	4
1	40	36	76	4	4
2	42	35	77	4	4
3	41	41	82	4	4
4	50	32	82	5	4
5	57	33	90	5	4
Total	293	199	492	26	24

Table VI.1.1. Sections required at the upper limit of 24, if the two dual language strands are combined on one campus. Sections required = combined enrollment ÷ class size, rounded up to the next whole section. At the upper limit of 24, the program needs 24 sections. Based on the 2025 distribution of 492 students; a 2027–28 plan should use projected rather than current enrollment (see V.5-F7).

CAUTION: The figures represented in the table above, although derived from the provided Class Size and Composition roster and are not guarantees that the student numbers and ratios will be preserved into consolidation.

RECOMMENDATIONS

OVERALL RECOMMENDATION

Build the consolidated program on a planning target of 20 with a ceiling of 24 students per section with at least two sections at every grade. Classroom language composition should not be set as a program-wide ratio but through documented grouping at each grade level due to natural shifts in EL status as students move through the program.

ID	PRIMARY RECOMMENDATIONS
VI.1-R1	Adopt twenty students as the planning target for every K–5 section, with the ceiling at 24. , Allocate sections to need rather than dividing enrollment evenly. Where <u>English learner</u> or special education density is highest, the sections should be smallest.
VI.1-R2	Run at least two sections at every grade, whatever the enrollment. A single-section grade cannot balance composition, has no partner for planning, and leaves very small numbers by the upper grades. If enrollment cannot support two sections at a grade, that is a finding about the program’s viability at that grade and should be reported as such.

ID	PRIMARY RECOMMENDATIONS
VI.1-R3	Revisit the Part 228 ceiling for dual language programs in light of forthcoming ISBE dual language guidance. The rule caps the bilingual ratio at 90 percent of the building general education ratio for the same grades.
VI.1-R4	Apply the one-third partner-language floor grade by grade. A merged program at 38 percent overall still breaches the floor at kindergarten and grade 5. Treat the kindergarten breach as real and the upper-grade breaches as not yet readable, because five years of <u>reclassification</u> have lowered EL counts in those grades while the same children remain enrolled.
VI.1-R5	Continue to develop a more nuanced language profile that follows each student. The program describes itself as 50/50, enrolls by a lottery meant to hold that balance. A home-language survey is the right first request but is an intake variable, collected once and static. What the program needs is a continuing measure of Spanish proficiency, which is what VI.3 specifies.
VI.1-R6	Make proficiency-balanced grouping a documented expectation rather than individual teacher design. Students interviewed describe it working if every grouping includes a strong Spanish speaker (model), a bilingual dictionary at each table (material support), and a shared script for what to say instead of “I don’t know” (language support). A composition target without a grouping expectation has only answered half of the composition and ratio question.
VI.1-R7	Set a ceiling of twenty-four students per section and say what happens when a section reaches it. A planning target guides how sections are built in the spring but does not cover a family that enrolls in October. The ceiling follows the same instructional logic as the target, beyond twenty-four a teacher cannot rotate groups of three to six through a thirty-minute daily block often enough for each group to be met, and it falls within both regulatory ceilings. Illinois sets no general class-size maximum.

SUPPORTING RECOMMENDATIONS

VI.1-R8 Answer the parent class size cap request Parents have asked the district to commit to a firm cap of 25 to 26 students or to a co-teaching model. **The program already runs well below that ceiling: the roster average is 18.9, no section exceeds 23, and observed attendance runs 14 to 19.**

VI.1-R9 Express the composition target as a decision rule, not a fixed ratio.

The applicant pool has already shifted unexpectedly: more Spanish-dominant than English-dominant applicants for the first time, while the kindergarten pipeline halves from six sections to a projected three. A rigid ratio calibrated to one year’s pool will be wrong by the time it is adopted. Instead, this chapter recommends a floor with a stated response: what the district does in a year when the pool skews heavily either way, who decides, and against what. Naming the response in advance gives the committee building sections in August a rule it can apply.

CAUTIONS ON USE

What these recommendations are and are not for. Design limitations of the study belong to Section VIII; this section addresses how this chapter is most likely to be misread.

CAUTION	WHAT TO STATE
SMALL CELLS	Once the program is at a single-site, Spanish counts will fall and some grade cells will become very small. The three Red Oak grades already fall below the one-third floor. They contain only five, five and seven students. Composition should not be published at a grade level below a stated minimum, and the district should set that minimum before the first report rather than after someone asks. Section VIII discusses the design limitation; this section sets the reporting rule.
CATEGORIES	Speaker groups are planning categories and should not be used as student labels. They exist to build sections, and they should be assigned by a written rule, dated, and revisited. Three points to consider: what happens when an individual student's group is genuinely ambiguous, which is common and increasingly so; in should be established that a simultaneous bilingual is a category in its own right rather than a rounding decision; and that the assignment is revisable both because children change and the district's own evidence says the population is becoming harder to sort.
USE AND NON-USE	Composition targets are for creating sections. They should not be used for admissions, individual placement, or staff evaluation. The evaluation has already found an English-language assessment functioning as an access gate to advanced coursework. A composition category could be misused the same way. The permitted and prohibited uses should appear in the same document that defines the fields.
FAMILIES	Composition ratios are the part of this report families are most likely to hear about second-hand, and the secondhand version will likely be simplified and less accurate. A family-facing explanation should exist in English and Spanish before the recommendations are adopted, and it should say plainly that a target describes how sections are built and never determines whether a particular child is admitted. The parent corpus already shows what happens without it: families in both languages describe the program's social life as skewing toward English-speaking families, a perception that a poorly explained ratio could reinforce.

SECTION VI · RECOMMENDATIONS BY DELIVERABLE (RFP SEC. D)

VI.2 · Curriculum Alignment, Coherence & Pacing Recommendations

Deliverable 2

DELIVERABLE	VI.2 · Curriculum Alignment & Pacing Recommendations
SCOPE	Recommendations for a unified K–5 curricular scope, sequence, pacing, and <u>language allocation</u> for the SD 112 Dual Language Program
ANCHORED BY	EQ 6 · Curriculum and Pacing
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillars: <u>Bilingualism</u> & <u>Biliteracy</u> , <u>Academic Achievement</u> , and <u>Sociocultural Competence</u> . Strand 2: Curriculum
RFP REFERENCE	RFP Sec. D

FINDING ID	WHAT IT ESTABLISHES
V.6-F1	Current delivered <u>language allocation</u> leans English-heavy against the stated 50% English/50% Spanish <u>program model</u> . This is evident in multiple data sources. <u>Language allocation</u> appears closer to 59%English/41% Spanish by instructional time and 65% English/35% Spanish across the entire school day; 60%English/40% Spanish by the literacy <u>pacing calendars</u> ' weekly minutes, and roughly 70%English/30% Spanish as teachers, parents, and students independently describe daily practice.
V.6-F2	5 of the 8 reviewed materials received the highest ratings for overall quality. Cross-cultural connections and <u>translanguaging</u> pedagogy scored lowest of seven areas reviewed on the materials evaluation. The curriculum materials each require more instructional time than the schedule allocates.
V.6-F3	Vertical coordination of the literacy calendar is solid. Grades 1 through 5 open and close all nine HMH modules within days of each other across the year, and the odd/even module-to-language-lead rule is applied identically at every grade band.
V.6-F4	The pacing plan has been designed to deliver two monolingual curricular across content areas. The district has already done some of the work to make connections across curricula, including a crosswalk document that connects core literacy and social studies, a <u>pacing calendar</u> for core literacy instruction and writing in English and Spanish, <u>pacing calendars</u> for Spanish Science and Social Studies, as well as a map of high priority foundational skills to teach in each language. Teachers have received professional learning sessions related to how to teach the K-3 HMH in the most efficient manner.

FINDING ID	WHAT IT ESTABLISHES
V.6-F9	The program has no defined K–5 progression for oral <u>language development</u> . Teachers report that there are no oracy <u>standards</u> for the dual language program; that is, no statement of what students should be able to produce orally, in each language, at each grade.

PREMISES

In dual language programs, in principle, the curricula in two languages are coordinated to build on one another. The district already has done some of the work to build cohesive plans between the two languages. The CKLA and Caminos foundational skills component of literacy instruction has a map to follow that prioritizes skills that are unique to each language. A curriculum crosswalk document exists that connects the HMH literacy modules to complementary social studies units. The recommendations in this section build on that existing work.

Designed language allocation and delivered allocation are different. The following features of the program in practice are different measures with different denominators: 59%English /41%Spanish instructional-time and 65%English /35 %Spanish whole-day (master-schedule analysis), 60%English /40%Spanish (pacing-calendar minutes), and approximately 70%English /30%Spanish (April 2026 observation, corroborated independently by teacher, parent, and student focus groups). The audit of the actual SY26 bell schedule totals 180 English/120 Spanish minutes per day (approx.60%English/40% Spanish). This section reports all figures, by source.

Dual language curricula in two languages are meant to converge with respect to instruction in multimodalities, translanguaging, and explicit cross-linguistic and cross-cultural transfer. In practice, this occurs at the word level, but can and should be extended across the other dimensions of academic language as well as modalities and cultural contexts.

ANCHOR FIGURE: PROPOSED DUAL LANGUAGE BILITERACY CURRICULAR FRAMEWORK

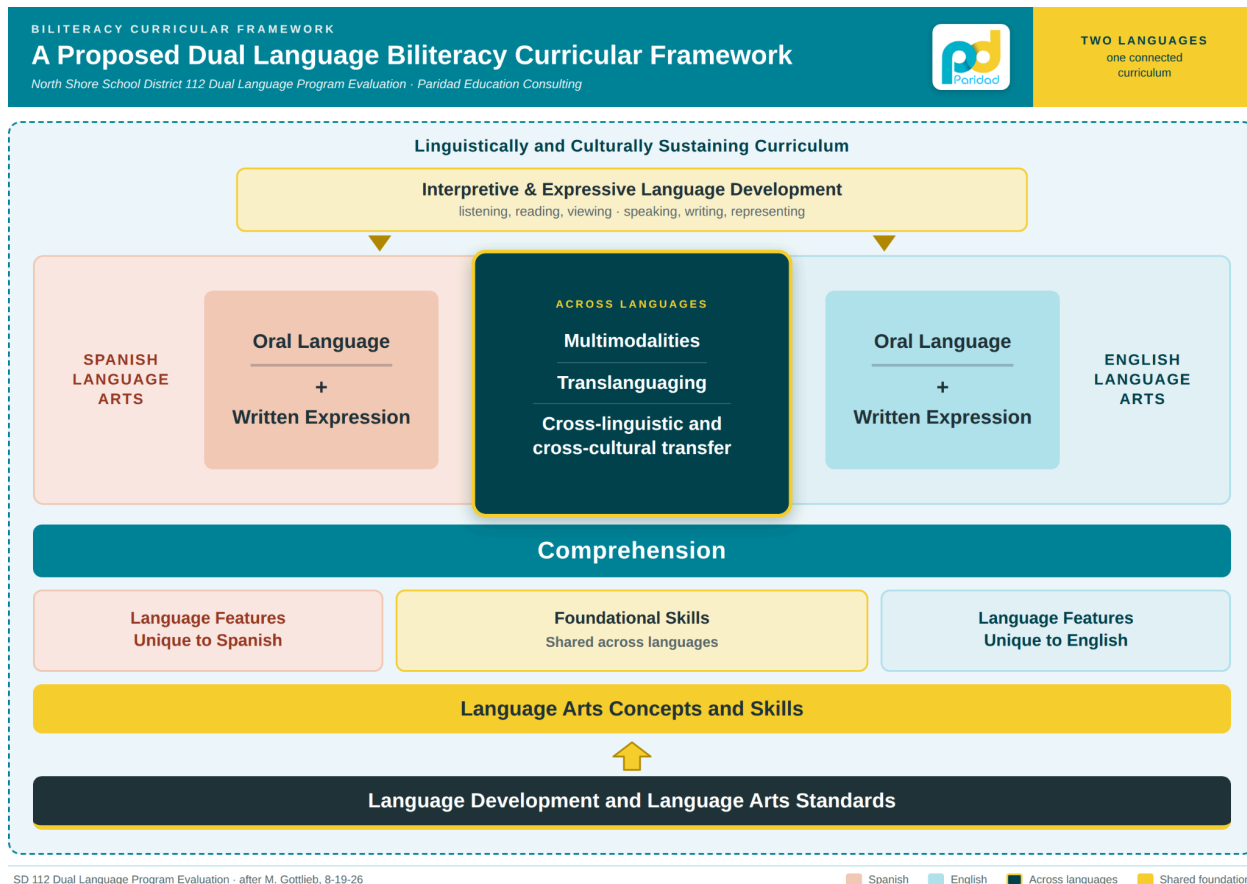


Figure VI.2.1. A Proposed Dual Language Biliteracy Curricular Framework: Example from language arts that illustrates components of a coordinated dual language curriculum framework.

This figure depicts a single coordinated biliteracy curriculum. Spanish Language Arts and language proficiency, and English Language Arts and language proficiency, each have their own oral-language, written-expression, and language-specific features, but both also rest on shared foundational skills and a shared comprehension layer, and both are designed to converge on multimodalities, translanguaging, and explicit cross-linguistic and cross-cultural transfer. The recommendations in this section of the evaluation report suggest that the same sort of mapping can be completed across all curricular areas: Spanish and English science, social studies and mathematics curricula. This figure represents how the components of language development and language arts can be mapped across two languages.

RECOMMENDATIONS

OVERALL RECOMMENDATION

Plan and implement the curriculum to increase the amount of instructional time in Spanish (especially Social Studies and Science) and intentionally point to the partner language during instruction in the other language throughout the content areas to move the program closer to the 50/50 language allocation model and build cohesion between the two languages throughout K-5.

ID	RECOMMENDATION
VI.2-R1	Align Pacing & Instruction: Develop a process for teachers, instructional coaches and administrators to work closely to protect the 50%/ 50% <u>language allocation</u> model making certain attention is paid to time every day for a science or social studies block (in Spanish) according to the published <u>pacing calendar</u> ;
VI.2-R2	Build in crosslinguistic connections: Identify and incorporate elements of crosslinguistic connections (for all dimensions of academic language) in all content areas regardless of the lead language for that week. Careful planning will allow students access to the <u>partner language</u> (across linguistic domains: listening, speaking, reading and writing) regardless of the lead language for that week.
VI.2-R3	Incorporate sociocultural contexts: Task the same or separate team to integrate examples of students' and families' funds of knowledge, as well as cross-cultural and sociocultural components across the curriculum. The <u>language arts</u> curriculum (HMH) makes authentic connections to students' backgrounds and languages in both the English and Spanish teacher editions. However, the CKLA Amplify Foundational Skills materials in English do not make connections to Spanish; the Spanish counterpart, <i>Caminos</i> foundational skills, does make reference and connections to English.
VI.2-R4	Follow published foundational skills scope & sequence: Maintain a boundary around the time allotted to CKLA/Caminos foundational skills to what the scope and sequence and <u>pacing calendars</u> recommend. 20-40 minutes a day total is sufficient. In practice, the time given for this component is sometimes 60-80 mins of instruction each day. This accounts in part for why time for the science and social studies blocks are reduced or limited.
VI.2-R5	Complete crosswalk between English and Spanish Curricula: Convene K–5 grade-level teacher teams to continue the crosswalk work that the Teaching & Learning department has begun. Conduct English/Spanish curriculum-alignment reviews, connecting the core literacy curriculum modules (grade level reading and writing in both languages) to science and social studies; identify complementary content that can be taught for transfer to the other language during <u>language arts</u> rather than teaching the exact same content twice in each language (e.g., reading complementary texts in each language rather than reading the exact same text within a module). This team or another can coordinate the content in the Spanish Number Corner with the English core mathematics unit topics.
VI.2- R6	Increase or add oracy components to instruction: Some curricular materials include ideas and strategies to increase oral <u>language development</u> throughout the modules and units of study; others do not. Grade level teams or discipline-specific teams can collaborate, review and make note of these strategies as well as add examples of peer to peer, small group, partner, large group interactions that can provide students opportunities to use their oral language skills in both languages every day.

SUPPORTING RECOMMENDATIONS

VI.2 - R7 Adopt a version of the Dual Language biliteracy curriculum framework (Figure VI.2.1) for the district's unified K–5 one coordinated bilingual curriculum mapping. This biliteracy curriculum framework connects directly to the proposed biliteracy assessment framework and can serve as a model for how to coordinate all the components of a content area across languages.

VI.2 - R8 Continue strong writing pedagogy and professional learning

Continue to build on the strong bilingual writing curriculum support given to teachers at grades 4 and 5 and extend writing support and coordination to grades K-3. Provide ample and consistent time each day for students to read complex, grade-level texts and write across the curriculum.

VI.2 - R9 Build on model practices

The evaluation team observed exemplary practices and knowledgeable staff at both schools. Developing a culture of peer-coaching and providing opportunities for practitioners to share their knowledge and skills with one another is an excellent resource within the program.

VI.2 - R10 Refresh cross-linguistic, translanguaging professional learning for all dual-language staff.

Provide refresher professional learning on translanguaging pedagogy, timed to support staff in preparing and implementing crosslinguistic connections across dimensions of academic language.

VI.2 - R11 Create a recurring mechanism for tracking designed-versus-delivered allocation.

The 70% English/30 % Spanish and 60% English/ 40% Spanish delivered figure is based on classroom-observations as well as teacher, parent and student descriptions of their lived experience. Recommend the district adopt a recurring, simple schedule audit or observation cycle (e.g., annual), so the gap between designed and delivered allocation can be tracked over time and addressed as needed.

VI.2 - R12 Professional Learning to support implementation

Provide guidance and professional learning for practitioners on implementing the existing CKLA and Caminos pacing calendar with examples of how to make crosslinguistic connections (across all dimensions of academic language) in each lesson; thereby keeping both languages connected regardless of the lead language for the week.

IMPLEMENTATION TIMELINE RECOMMENDATIONS

The graphic below suggests an implementation map beginning in the 2027-2028 school year.

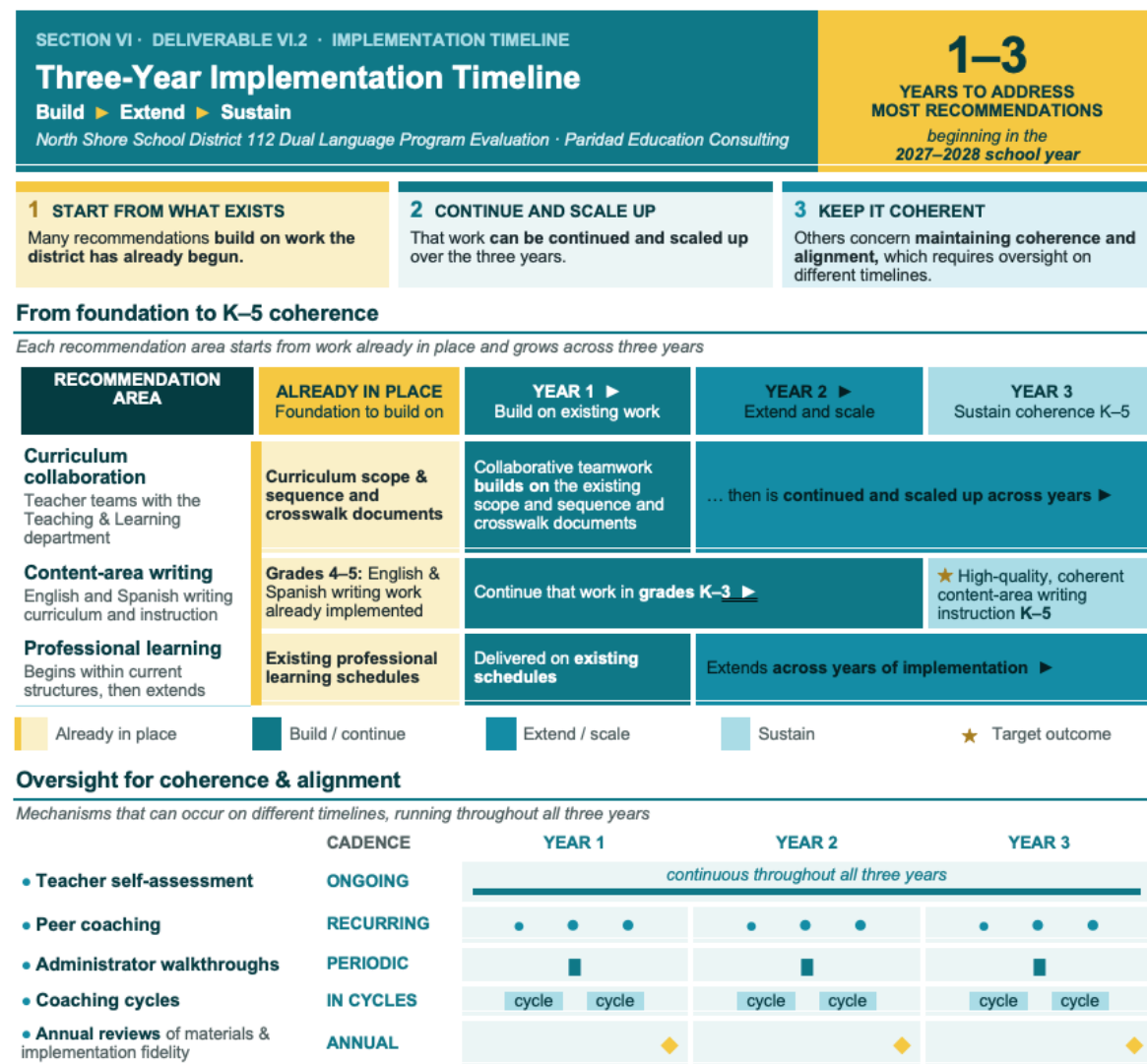


Figure VI.2.2. Provides a Suggested Implementation Map for the Recommendations in this Deliverable

CAUTIONS ON USE

The cautions below address how these recommendations could be misread. Design limitations of the study are discussed in Section VIII.

CAUTION	WHAT TO STATE
USE AND NON-USE	These pacing guide recommendations are meant to support coherence and are not a compliance tool. It should not be used for individual teacher evaluation or to judge a single classroom's progress in a given week against the calendar.
DESIGNED VS. DELIVERED	Every allocation figure in this chapter states whether it is a designed figure (master schedule, <u>pacing calendar</u>) or a delivered figure (classroom observation, stakeholder report), and by what method it was established. Reporting the two interchangeably (for example, citing the corrected 59/41 (or 65/35) master-schedule figure as if it described daily classroom practice) will mislead, given that independent measures put delivered allocation closer to 70/30.



CAUTION	WHAT TO STATE
ACCESSIBILITY	These pacing guide recommendations assume a stable roster present for the full instructional sequence. It does not, by itself, address what happens for a newcomer entering mid-year and mid-module, or for a student with an IEP whose service minutes cut across the allocation

SECTION VI · RECOMMENDATIONS BY DELIVERABLE (RFP SEC. D)

VI.3 · Biliteracy Assessment Framework Recommendations

Deliverable 3

DELIVERABLE	VI.3 · <u>Biliteracy</u> Assessment Framework Recommendations
SCOPE	A balanced <u>biliteracy</u> assessment system measuring literacy and oracy in both languages, K–5.
ANCHORED BY	V.4: EQ 4 · <u>Biliteracy</u> Assessment
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	Pillar: <u>Bilingualism</u> and <u>Biliteracy</u> . Strand 4: Assessment and Accountability
RFP REFERENCE	RFP Sec. D

Recommendations for a comprehensive biliteracy assessment framework, one of the deliverables required in the North Shore SD 112 RFP, has been conceptualized from research-informed, standards-referenced evidence in addition to qualitative and quantitative data sources from this evaluation. It draws from the following evaluation questions: EQ 1 (the extent of English language proficiency for ELs), EQ 2 (the extent of Spanish language proficiency for all dual language students), and EQ 3 (academic achievement of ELs). It directly addresses EQ4 (the district’s instruments and measures in the dual language education program).

FINDING ID	WHAT IT ESTABLISHES
V.4-F1	English <u>language proficiency</u> is measured annually across all four <u>language domains</u> (listening, speaking, reading, and writing) by <u>WIDA ACCESS for ELLs</u> , with the <u>WIDA Screener</u> at entry. A Spanish <u>language proficiency</u> measure is not in use.
V.4-F2	Spanish <u>language arts</u> reading and foundational literacy are broadly comparable to English from kindergarten - grade 5. <u>NWEA MAP</u> Lectura has not been normed on English-speaking students.
V.4-F3	Spanish and English writing are collected four times a year at grades 4–5 and to our knowledge, interpreted using the WIDA Language Charts
V.4-F4	The relationship between the two languages is not captured; this includes cross-linguistic or cross-cultural transfer, <u>translanguaging</u> , or multimodal expression.
V.4-F5	Spanish <u>academic achievement</u> is measured in reading only.

This section of the deliverable presents the evaluation's findings that impact the design and recommendations for a comprehensive biliteracy assessment framework for North Shore SD 112. It is essentially an assessment audit based on an inventory of the district's current measures.

PREMISES

Conceptualization of the Comprehensive **Biliteracy** Assessment Framework

The suggested comprehensive biliteracy assessment framework has been built from an evidence-based biliteracy curricular framework informed by the three pillars of dual language education, shown below. Here the program's two languages, Spanish and English, develop and converge into a single system. In it, oral language and written expression for English and Spanish language development, derived from language development standards, are connected through cross-linguistic and cross-cultural transfer, multimodalities, and translanguaging. At the same time, Spanish and English language arts' concepts and skills, rooted in the state's academic content standards, blend together, generally at the higher ends of language proficiency.

Comprehension is an outgrowth of the two languages (Spanish and English) and the two constructs (language development and language arts). Here the comprehensive biliteracy curricular framework is applied within a linguistic and culturally-sustaining curriculum for all students in the K-5 dual language program. There is recognition of the unique features of each language while there is a convergence of the foundational skills shared across the two languages.

In its totality, the comprehensive biliteracy curricular framework is implemented within a balanced linguistic and culturally sustaining curriculum for all students in the dual language program; the biliteracy assessment framework is designed to reinforce the curriculum shared throughout the district.

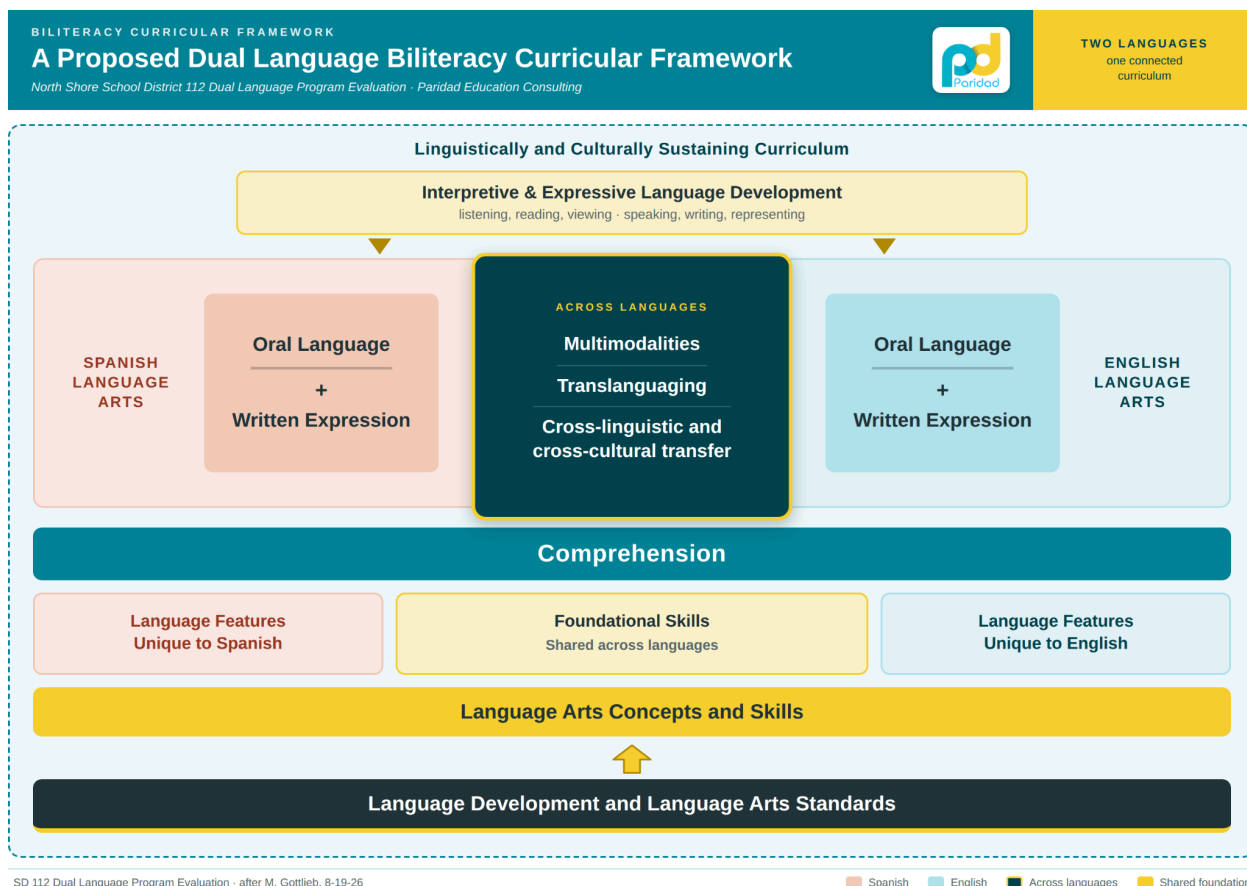


Figure VI.3.1. A Proposed Dual Language Biliteracy Curricular Framework onto which the Biliteracy Assessment Framework is Superimposed.

THE ANCHOR FIGURE

The proposed comprehensive biliteracy assessment framework shown below is superimposed onto the curricular framework to create a 1:1 correspondence between teaching and learning in the dual language program. Together, they create a balanced and aligned system representing the assessment of language proficiency and language arts as a reflection of curriculum design. Mathematics and other content areas, such as social studies and science, can be easily added to the framework and implemented under the same conditions. In the biliteracy assessment framework, there is a proposed measure for each component based on existing ones— in Spanish, in English, and across the two languages. Elements shown in red are not currently in place in North Shore SD 112 and are suggested in the recommendations to strengthen biliteracy and bring coherence to the system.

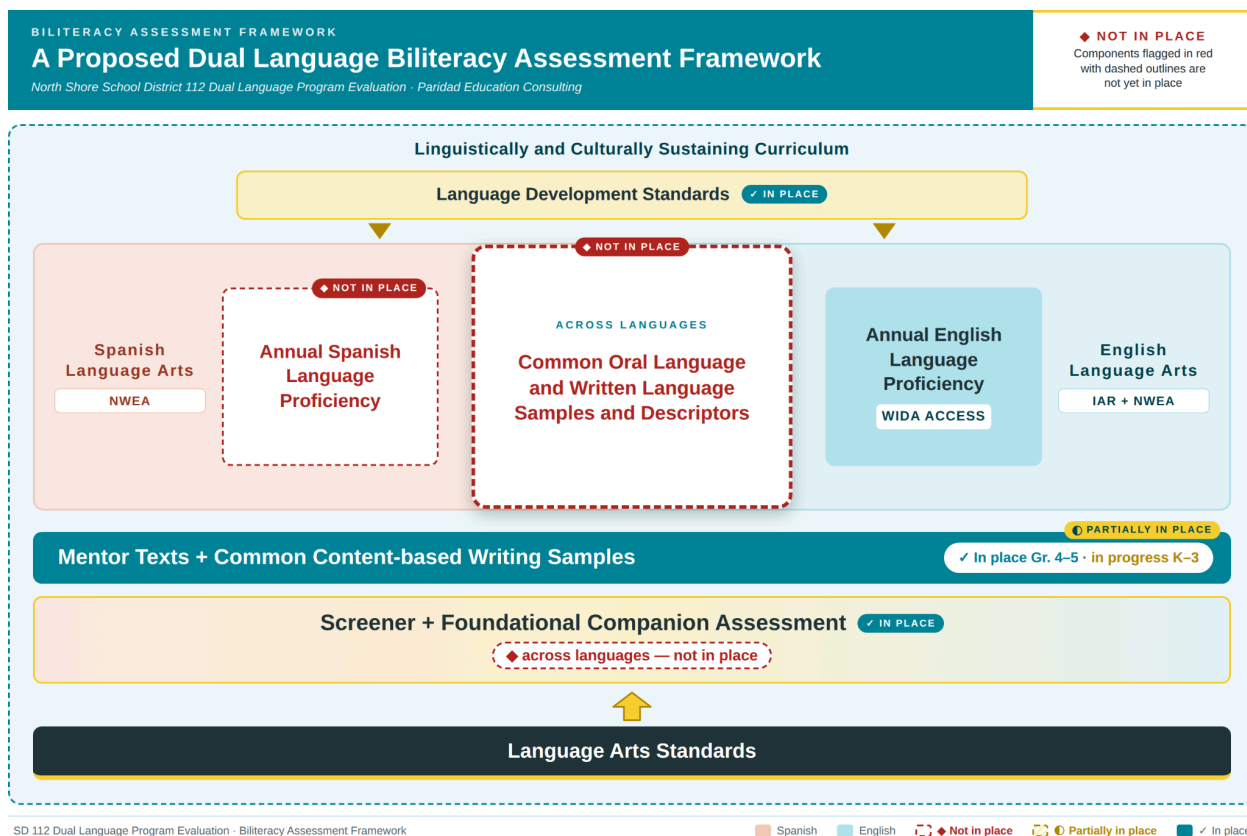


Figure VI.3.2. A Proposed Dual Language Biliteracy Assessment Framework. Red font indicates measures not currently in place.

RECOMMENDATIONS

Recommendations for the Comprehensive Biliteracy Assessment Framework as a Whole

OVERALL RECOMMENDATION

Construct and maintain a comprehensive biliteracy assessment framework in tandem with and complementary to that present in the general education assessment, where appropriate. Ensure that the biliteracy assessment framework is balanced across languages (Spanish and English) and constructs (language development and language arts) for all students in the dual language education program.

Supporting Recommendations

The findings from the analysis of the instruments and measures point to a series of supporting recommendations that are intended to strengthen assessment within the dual language education program, taking the existing measures into account. Together the recommendations merge to create a balanced dual language biliteracy assessment framework reflective of the three pillars of dual language education.

VI.3.R1 Design and adopt a balanced biliteracy assessment framework organized by purpose and language.

The evaluation team recommends that the district organize assessment around three purposes:

1. initial screening and collection of baseline data
2. documentation of interim progress within and across language arts and language proficiency, and
3. determination of annual growth in both English and Spanish, as elaborated in the suggested biliteracy assessment framework calendar.

In addition, as noted in the table below, the team suggests that the framework also addresses the connections between the languages at the word/phrase, sentence, as well as discourse dimensions. All assessments should point to one of the three dual language assessment pillars.

Purpose for Assessment	Recommended in English	Recommended for cross-linguistic & cross-cultural connections	Recommended in Spanish
1. Screen / identify <u>English learners</u> (at entry) and collect baseline data for oral language (<u>bilingualism</u>) and <u>biliteracy</u>	WIDA <u>language proficiency Screener</u> + brief oral language sample and content-based writing sample Literacy <u>screener</u> (if district mandated)	Initial content-based oral language and written language samples	Home-language survey for <u>English learners</u> + brief oral language sample and content-based writing sample for all dual language students
2. Monitor progress (interim / benchmark) in <u>Language Arts</u> and <u>Language Proficiency</u>	<u>NWEA MAP</u> Growth (reading); curriculum-based oral and written <u>language proficiency</u> tasks interpreted such as with the WIDA Language Charts	Oral language and writing samples incorporating multimodalities to encourage <u>translanguaging</u> , and cross-linguistic & cross-cultural transfer	<u>MAP Lectura</u> for <u>English learners</u> (if useful); Spanish common oral language and written language samples systematically collected for all dual language students
3. Measure growth in English (for annual accountability) and Spanish <u>language proficiency</u> (on an interim basis)	<u>WIDA ACCESS</u> ; WIDA MODEL- optional interim test; oral language and writing samples uniformly interpreted, such as with the WIDA Language Charts		Uniformly interpreted common oral language and writing samples using a standards-referenced tool, such as the WIDA <u>DALE Proficiency Level Descriptors</u>
Measure annual growth in English and Spanish <u>Language Arts</u> (for accountability)	<u>IAR</u> : English <u>language arts</u>		A Spanish <u>Language Arts</u> test normed on both Spanish and English speakers

Table VI.3.3 Suggested Pool of Assessment Instruments by Purpose and Language

VI.3.R2 Design and adopt common measures to document language growth and illustrate the confluence of languages.

The highest-priority additions to the current biliteracy assessment framework are on the Spanish side and on cross-linguistic connections. The evaluation team recommends that the district adopt: (a) a common Spanish oral language-sample protocol; (b) a common Spanish writing

protocol; (c) K–2 Spanish and English foundational literacy benchmarks; and (d) a cross-linguistic transfer inventory. Oral language and original written student samples could be interpreted by grade-level teacher teams so results mean the same across classrooms and grade levels. Language development protocols can be adopted or adapted from the WIDA Language Charts (2025) and WIDA Los descriptores del nivel de desempeño lingüístico (2023).

VI.3.R3. Disaggregate language development data by learner language category.

To the extent feasible, the evaluation team recommends reporting Spanish language development (and cross-linguistic) growth data separately for primary English speakers, heritage Spanish speakers, and primary Spanish speakers (English learners). All groups should show steady progress primary-Spanish students should be advancing at a faster rate in Spanish as they move toward more complex literary Spanish. Given this information, more targeted instruction can be planned.

VI.3.R4. Use technology to augment documentation and archive assessment data.

Teachers already appear to be overburdened by English assessment, leaving little time for comparable assessment in Spanish. The evaluation team recommends using recording plus automatic transcription (e.g., a district-approved, student-safe tool) to capture and analyze English and Spanish oral language and writing samples within Professional Learning Communities or grade-level teams. Agreement in teacher interpretation using reliable AI tools plus teacher validation will yield systematic evidence of Spanish and English oracy and writing.

VI.3.R5. Create a biliteracy portrait for every dual language student.

The evaluation team recommends replacing per-instrument, per-language reporting with an individual student assessment portrait that places Spanish, English, and their cross-linguistic connections side by side over the course of an academic year. In that way, a student's individual growth in biliteracy, within and across Spanish and English, will be visible to students, teachers, families, and leaders. Initially, student portraits could be co-constructed between students and teachers on a trial basis and embedded in curriculum; once there is agreement on the entries, the portraits can be digitally stored and students can have some choice in their selections.

VI.3.R6. Revisit the biliteracy assessment framework and evaluate its effectiveness on an annual basis.

Annual surveys collected from administrators, teachers, families, and students on the perceived comprehensiveness and effectiveness of the biliteracy assessment framework should be administered that yield feedback to be used for continuous improvement of the curricular and assessment frameworks and the dual language education program.

Summary of the scope of the comprehensive biliteracy assessment framework

The comprehensive biliteracy assessment framework as conceptualized should be fully integrated into SD 112's overall assessment system. It is designed to enable dual language students and educators to have access to:

- Academic content standards for English and Spanish language arts to measure grade-level literacy concepts and skills, minimally with shared foundational skills and comprehension strategies with consideration of other academic content standards, such as mathematics, science, and social studies

- English and Spanish language development standards to document student growth in language development across levels of language proficiency and four language domains — listening, reading, and viewing (interpretive language) and speaking, writing, and representing (expressive language)
- A blending of Spanish and English language development and Spanish and English language arts skills and concepts exemplified through translanguaging, multimodalities, and cross-linguistic transfer
- A linguistic and culturally sustainable biliteracy curriculum across the content areas (mathematics, science, social studies)

The figure below summarizes the areas of assessment that are currently in place and outlines the types of instruments and measures the district should consider per their purposes and use cases.

VI.3.R7: It is suggested that the central data archive be enhanced to house and systematically review both qualitative and quantitative assessment data in Spanish and English across the three pillars of dual language education for all students in the district’s dual language education program.

Area of assessment	Currently in place	Suggested Recommendations
English <u>language proficiency</u> (listening, speaking, reading, writing, literacy, comprehension)	<u>WIDA Screener</u> ; <u>WIDA ACCESS</u> : the annual ELP assessment	Maintain an archive of <u>WIDA Screener</u> and annual <u>WIDA ACCESS</u> data
Early (K–2) English foundational skills		
English <u>language arts</u>	<u>IAR</u> ; <u>NWEA MAP</u>	Maintain an archive of <u>IAR</u> and <u>MAP</u> ELA data
Spanish <u>Language Arts</u> (lectura)	<u>NWEA MAP</u> : Lectura	Reconsider value of SLA interim standardized test data for English speakers
Spanish oral language & writing proficiency and achievement		Create a common assessment to measure on an interim basis or seamlessly incorporate into curriculum; adopt/adapt a uniform <u>rubric</u> and establish inter-rater reliability among teachers
Early (K–2) Spanish foundational skills		Consider overlap with English foundational skills to create <u>cross-linguistic transfer</u>
Cross-linguistic and cross-cultural transfer, an expression of socio-cultural competence		Establish a <u>translanguaging</u> policy, curricular space within the pacing guide, and a mechanism to capture students’ blended use of languages, ideally through common assessment

Table VI.3.4. Areas of Assessment Currently in Place and Suggested instruments

Overall conclusions from the findings indicate that measures in English predominate and there is greater reliance on standardized, normed reference tests than classroom-based measures. Presently there is an array of measures, mostly in English although some in Spanish, however, they do not produce a coherent biliteracy assessment framework.

The following set of recommendations, based on the findings from data analysis point to how to strengthen the current biliteracy assessment framework.



Paridad Education Consulting · SD 112 DL Program Evaluation · Spanish Assessment Instruments K-5

Figure VI.3.5. Considerations for Annual and Interim Instruments and Measures for Spanish language Proficiency, Spanish Language Arts, and Mathematics

Implementing the Recommendations

The biliteracy assessment framework, as conceptualized, represents a balanced system of languages (Spanish and English) and constructs (language development and language arts). The confluence of these languages and constructs allows for cross-linguistic transfer through translanguaging coupled with multimodal representation. Students' common oral language and written language samples collected, interpreted, and systematically archived give a full and authentic picture of how students in the dual language learn and how their teachers teach.

The recommendations for such a biliteracy assessment framework can be phased in over the next couple of years alongside the building of the 50–50 instructional model: design and pilot common Spanish oral language and writing samples and their interpretation first, then consider building the full system and calendar, including standardized annual (and interim) measures. The evaluation team recommends allocating protected teacher time for interpreting and consensus building on the oral language and written language samples within grade level teams, existing PLCs, and professional development.

CAUTIONS ON USE

Recommendations carry some cautions in their implementation; here are some to take into account in developing the biliteracy assessment framework.

CAUTION	WHAT THE EVALUATION TEAM STATES
USE AND NON-USE	These measures are to be evaluated in their totality; that is, their impact as a balanced, coherent system with state, district, and classroom contributions to the dual language education program. They are not intended for teacher evaluation. Teachers should understand the purpose and value of each measure before the first team-interpretation session.
COMPARABILITY	An <u>ACCESS proficiency level</u> , a <u>MAP RIT</u> , and a locally interpreted 1–4 <u>rubric</u> are not on a common scale and cannot be averaged or plotted together. Each language's growth should be tracked against its own trajectory.
LOCAL MEASURES	Locally interpreted samples are trustworthy when they include common prompts, a uniform <u>rubric</u> , calibrated team interpretation, double-scoring of a subset of samples, and a periodic drift check, and the inter-rater reliability.

SECTION VI · RECOMMENDATIONS BY DELIVERABLE (RFP SEC. D)

VI.4 · Program Model Recommendations

Deliverable 4

DELIVERABLE	VI.4 · <u>Program Model</u> Recommendations
EVALUATION QUESTION	<i>EQ 7: How effective is the current 50/50 model, and what model is recommended going forward?</i>
IN OPERATIONAL TERMS	<i>What does the combined evidence show, and how does it point to the most appropriate dual language <u>program model</u> for <u>English learners</u> in District 112?</i>
SCOPE	The instructional model recommendation, 50/50 as currently delivered or a transition to another model.
ANCHORED BY	All six EQs Data Analyses
DUAL LANGUAGE PILLARS & GUIDING PRINCIPLES STRANDS	All <u>three pillars</u> and all seven <u>guiding principles</u> and corresponding strands.
RFP REFERENCE	RFP Sec. D

THE PROGRAM-MODEL QUESTION: EQ 7 ANALYSIS

FINDINGS SUMMARY

District 112 has not been running the 50/50 model that it asked to be evaluated. In practice, the school day is closer to 60% English/ 40% Spanish. Revisions to time allocation need to be made to conform with a true 50/50 model along with inclusion of appropriate Spanish language measures before an assessment of this model can be validly determined

DATA SOURCES

QUANTITATIVE	QUALITATIVE
All the quantitative data from EQs 1-6	All the qualitative data from EQs 1-6

METHODOLOGY

Data and findings from six chapters were reviewed and analyzed as they applied to the question of the effectiveness of SD 112's current 50% English/ 50% Spanish language allocation model. The analysis followed the three steps discussed below:

- 1.) Establish what language allocation is actually being delivered before asking whether it works. Allocation evidence (V.6-F1, V.2-F2) is read first and independently of any outcome
- 2.) Read outcomes separately by language and by pillar (Bilingualism/ Biliteracy, High Academic Achievement, and Cultural Competence): English evidence (V.1, V.3-F1, V.3-F2) and Spanish evidence (V.3-F3, V.2-F1)
- 3.) Weigh the convergence of the evidence to make a claim. Regardless of what chapter the findings come from, the convergence of four or more methods will establish a strong finding and claim.
- 4.) Ask what the program measures on both the English side and the Spanish side. English proficiency is measured annually across four domains; Spanish proficiency is not measured at any grade, and Spanish oral language is not measured at all (V.4-F1). The bilingualism/ biliteracy pillar is not completely measured at this time.
- 5.) Ask what the program measures in terms of Sociocultural competence

KEY FINDINGS

FINDING ID	WHAT IT ESTABLISHES	ROLE IN THE ARGUMENT
WHAT THE PROGRAM DELIVERS		
V.6-F1	Delivered <u>language allocation</u> leans English-heavy against the stated 50% English/50% Spanish <u>program model</u> . <u>Language allocation</u> appears closer to 59% English/41% Spanish by instructional time and 65% English/35% Spanish across the entire school day, and roughly 70% English/30% Spanish as teachers, parents, and students independently describe daily practice. The literacy <u>pacing calendars</u> arrive at 50% English/50% Spanish by the literacy <u>pacing calendars</u> ' weekly minutes.	Four measures of delivered allocation, each with its own denominator, and none reaches 50/50.
V.2-F2	Seven independent sources agree that daily <u>language allocation</u> runs substantially more English-heavy than in the stated 50% English/50% Spanish design, closer to 70% English/30% Spanish in instructional practice according to students, teachers and parents.	Corroborates the practice-level figure from observation, schedule, and stakeholder report.
V.6-F2	5 of the 8 reviewed materials received the highest rating for overall quality. Cross-cultural connections and <u>translanguaging</u> pedagogy score lowest of seven areas reviewed on the materials evaluation. The curriculum materials each require more instructional time than the schedule allocates.	The materials are rated highly; the constraint is instructional time. Bounds any claim that the literacy core delivers its full design.
V.3-F9	Mathematics is taught in two languages and assessed in one. The daily program pairs a 60-minute Math block in English with a 30-minute Number Corner block in Spanish at every grade, while both <u>IAR</u> and <u>NWEA</u> mathematics are administered in English.	Locates part of the English lean in the content blocks.

KEY FINDINGS		
FINDING ID	WHAT IT ESTABLISHES	ROLE IN THE ARGUMENT
V.6-F6	Science and social studies in practice are the first areas of instruction cut under time pressure, and they represent an important block in the program's Spanish side of instruction.	Spanish content time is what gets cut, which supports the exposure argument.
V.1-F15	Classroom observation showed that English is prominent across language blocks. In Spanish-designated blocks students frequently answered in English and teachers accommodated by translating rather than redirecting.	Shows the lean towards English also occurs inside Spanish-designated time, not only outside the literacy core.
ENGLISH OUTCOMES		
V.1-F1	<u>English Learners</u> in the two dual language schools made large, sustained composite growth across three springs.	English proficiency growth (<u>English Learners</u> only).
V.1-F2	Growth was balanced across oral language and literacy.	Growth is not confined to one domain.
V.1-F3	Benchmarked against WIDA Consortium averages, three of the four grade <u>cohorts</u> met or exceeded the matching consortium gain across the two-year window while the oldest <u>cohort</u> finished modestly.	The benchmark behind the growth claim, with its exception.
V.1-F5	By 2024–25, 19.3 percent of <u>English Learners</u> had reached the Illinois 4.8 <u>reclassification</u> criterion (an estimate pending confirmation against district <u>reclassification</u> records).	<u>Reclassification</u> status
V.3-F1	Dual language students grew faster than national peers who began at the same score in English, in all three years and in both English <u>language arts</u> and mathematics in English.	Academic growth in English for the whole program population.
SPANISH OUTCOMES		
V.3-F3	Dual language students grew more slowly than national peers who began at the same score in all three years in Spanish <u>language arts</u> (reading).	Spanish reading growth was below national <u>norms</u> in each of the three years.
V.3-F8	Teachers describe Spanish reading and Spanish <u>academic achievement</u> as falling short while English outcomes continue to grow.	Stakeholder corroboration of the Spanish growth pattern.
V.2-F1	Sampled Spanish writing (n = 4) are at an intermediate level, as interpreted by two independent scoring <u>rubrics</u> , the district's SLA <u>rubrics</u> and WIDA's <u>MARCO DALE</u> descriptors, yielding an identical rank order across both.	Students' Spanish writing results need more data over time to determine progress in this <u>language domain</u> .
WHAT THE DISTRICT CAN MEASURE		
V.3-F4	The district's only standardized test in Spanish is for Spanish <u>Language Arts</u> .	Spanish outcome evidence currently rests on one instrument.
V.4-F1	English <u>language proficiency</u> is measured annually across all four <u>language domains</u> (listening, speaking, reading & writing) by <u>WIDA ACCESS</u> , with the <u>WIDA Screener</u> at entry. No Spanish <u>language proficiency</u> measure is in use.	Only the English half of the <u>bilingualism</u> pillar is measured.

KEY FINDINGS		
FINDING ID	WHAT IT ESTABLISHES	ROLE IN THE ARGUMENT
V.4-F5	Spanish <u>academic achievement</u> is measurable in <u>language arts</u> (reading only). <u>NWEA</u> Mathematics in Spanish exists; however, mathematics is generally taught in English.	A Spanish mathematics measure exists but is not administered.
V.4-F4	No documentation of the relationship between the two languages: cross-linguistic or cross-cultural transfer, <u>translanguaging</u> , or multimodal expression.	No current measure captures what a student can do across both languages.
WHAT THE BUILDING AND COMPOSITION PERMIT		
V.5-F4	The shift to lower numbers of EL Spanish speakers from Kindergarten to grade 5 is a result of ELs students who are <u>reclassified</u> out of EL status and stay in the dual language classroom, so the count on paper falls while the actual makeup of the room is unchanged.	Which composition counts are accurate? Kindergarten is more accurate. The upper grades understate Spanish speakers because of the reclassification of ELs who are still in the classrooms.
V.5-F8	In a dual language program, <u>reclassification</u> changes only the student's label. The student keeps the same seat in the same section, receives the same instruction in the same two languages, and remains exactly as Spanish-speaking as the day before. The program's single composition variable therefore declines across the grade span because of how students are labeled, and fastest in the upper grades.	Any grade 5 composition shortfall is at least partly an artifact of the label.
V.5-F5	The district continues to augment its language screening at intake to more specifically account for simultaneous bilingual learners and <u>heritage speakers</u> , so the composition half of EQ 5 is more complex than the question suggests.	Composition is more than a two-way ratio; the basis for the linguistic portrait in VI.4-R10.
V.5-F7	The 2025 roster is not the correct planning basis as it doesn't account for projections that have student enrollment fall to 447 by 2026–27 school year.	Sets the planning base for the merged program under either model.
V.5-F6	Illinois rule caps the bilingual student-teacher ratio at 90 percent of the building general education ratio for the same grades. It is the only legally binding figure among those considered.	A fixed constraint under 50/50 and 80/20 alike.

PREMISES

- **The model recommendation draws on the whole evidence base.** No chapter in Section V answers the Program Model question on its own. The convergence of the data lead to the recommendations. V.6 demonstrates the allocation evidence, V.3, V.2, and V.1 show the language proficiency and academic achievement outcomes, V.4 what is measured and what is not measured, and V.5 describes what the building and the composition will permit.
- **Because the 50/50 model is not being fully delivered, it cannot be tested directly.** The evaluation can report what happened in days running near 70% English/30% Spanish in practice and from independent reporting from teachers, students and parents. The remaining data collected was from a language allocation closer to %60 English / 40% Spanish. This chapter's statements about the current model describe delivered practice.

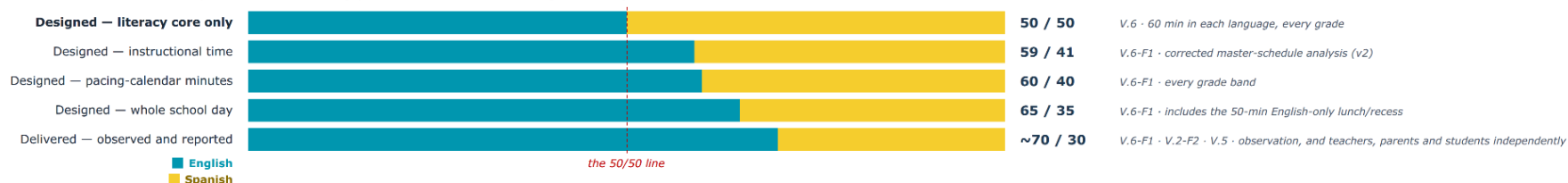
THE ANCHOR FIGURE

FIGURE VI.4 · THE PROGRAM-MODEL RECOMMENDATION

Evidence from V.1–V.6 against the three pillars, and what each model would change · SD 112 Dual Language Program Evaluation

1 WHAT IS ACTUALLY BEING DELIVERED

The stated model is 50/50. No measurement point reaches it, and the gap widens the closer the measure gets to lived practice.



The English lean originates entirely outside the literacy core — Math and CAPE in English, a standing 60-min/day English advantage, plus lunch and recess.

2 THREE PILLARS: ASSESSMENT BY LANGUAGE

Measures as confirmed in Figure V.4. An empty box is a finding, not a blank.

□ In use districtwide □ In use at teachers' discretion □ Not in place Teal text: what the evidence shows

	IN ENGLISH	IN SPANISH	ACROSS THE TWO LANGUAGES
Bilingualism & Biliteracy	WIDA ACCESS, four domains, annual — English learners only · WIDA Screener at entry MAP Reading · IAR ELA (gr 3–5) · FastBridge · Into Reading · CKLA (K–2) Interim English proficiency measure not in place. V.4-F6 V.1 +76 ACCESS scale-score points across three springs (n = 119); 19.3% reached the 4.8 criterion V.1-F5	MAP Lectura K–5 (not normed on English speakers) · FastBridge Español iArriba la Lectura! · Caminos (K–2) · gr 4–5 writing sample at teachers' discretion Spanish proficiency measure or oral measure not in place. V.4-F1 · V.4-F6 V.3-F3 growth 31.1 / 36.0 / 36.5 against a norm of 50; standing falls from about the 40th to the 26th percentile; grade 5 at the 20th	Instrument assessing what a student does across the two languages not in place. V.4-F4
Academic Achievement	MAP Mathematics · IAR Mathematics (gr 3–5) · Bridges 9 of 10 V.3-F1 conditional growth above the 50th norm in 6 of 6 subject-years; status at the 60th–71st percentile	Science and social studies at teachers' discretion only: Social Studies Alive! 8 of 11 · FOSS 3 of 10 Mathematics measure or content assessment. NWEA Mathematics in Spanish exists and is not in use. V.3-F4 · V.4-F5	One-third of designed mathematics time (Number Corner, in Spanish) is not assessed on any standardized measure. V.3-F9
Sociocultural Competence	Sociocultural competence measure in English not in place.	Sociocultural competence measure in Spanish not in place. V.2-F8 students' orientation toward bilingualism is positive — reported, not measured	Measure of bicultural or cross-cultural competence not in place at any grade. V.4-F4

Nothing is in place in the across-languages column in any row, no sociocultural measure is in use in either language, and outside reading and early literacy Spanish has no districtwide measure. The comparison a dual language program most needs cannot be made at all.

3 MODELS CONSIDERED — WHAT EACH WOULD CHANGE

50/50 CONTINUED, AS DELIVERED

Change nothing.

- Keeps a 70/30 English day under a 50/50 name.
- Spanish standing continues its three-year decline.
- It is not the stated model.

NOT RECOMMENDED

50/50 DELIVERED AS SPECIFIED

Repair the four fixed-language blocks and lunch/recess.

- Recovers ~9–15 points of Spanish time at no instructional cost.
- Tests the stated model for the first time.
- Does not add Spanish minutes beyond the design.

RECOMMENDED

Figure VI.4.1. The program-model recommendation: evidence from V.1–V.6 against the three pillars, and what each model would change.

RECOMMENDATIONS

OVERALL RECOMMENDATION

Adapt the current district instructional and assessment frameworks to attain a full 50/50 dual language education model. To achieve this goal, the district needs to implement a curricular framework that specifies a coordinated delivery of Spanish and English instruction for both language development and academic achievement. In parallel, the district needs to implement a comprehensive biliteracy assessment framework that reflects the dual language curriculum. Together both frameworks should provide space to incorporate cross-linguistic and cross-cultural considerations.

ID	PRIMARY RECOMMENDATION
VI.4-R1	Adopt 'deliver 50/50 as specified' as the district's <u>program model</u> position.
VI.4-R2	Re-adjust the four fixed-language content and the specials blocks so that Spanish time reaches the designed allocation and more opportunities in Spanish in non-instructional settings.
VI.4-R3	Evaluate and publish an instructional allocation figure each year, by grade and by block, against the designed percentage.
VI.4-R4	Protect the English <u>language proficiency</u> and achievement results explicitly through to the full implementation to the 50/50 model. Treat <u>ACCESS</u> matched-cohort growth and the <u>MAP</u> conditional-growth series as the baseline and review annually as a way to monitor.
VI.4-R5	Adopt a K-5 Spanish <u>language proficiency</u> measure of record, covering listening, speaking, reading and writing, and administer
VI.4-R6	If the use of <u>MAP Lectura</u> continues, then it should only be administered to native speakers of Spanish. According to the <u>NWEA</u> technical manual, Non-native speakers of Spanish are not part of the norming population.
VI.4-R7	Administer <u>NWEA</u> Mathematics in Spanish only once the program has achieved the true 50/50 allocation; presently most mathematics instruction, other than numbers corner, is in English.
VI.4-R8	Coordinate English and Spanish curricular materials and <u>translanguaging</u> capacity. Building coherence across the two sets of curricula, builds efficiency and more minutes of instruction in each language and across the two languages
VI.4-R9	Collect and create a complete linguistic portrait that includes both initial language(s) upon entry into the program, bilingual development and plan the merged program's composition grade by grade rather than on the program-wide average, which hides a shortfall below the required minimum at kindergarten and grade 5.

VI.4-R10

Build the review point into the calendar. Beginning in the 2027-2028 school year, build in a time for an annual review of the language allocation to see progress toward reaching a true 50% English / 50% Spanish language allocation.

IMPLEMENTATION RECOMMENDATIONS

Some of the elements that should be in place to facilitate implementations in this deliverable:

- rebuild the schedule so that delivered allocation reaches the designed 50%English/ 50% Spanish, figure;
- name and pilot a Spanish proficiency measure;
- begin to assess mathematics bilingually;
- use the augmented bilingual learner portraits to follow the progress of the wide range of bilingual learners in the dual language program.

Starting in the 2027-2028 school year, as the district begins to implement the recommendations from this report, the dual language program will gather bilingual data: two-three years of delivered-allocation reporting, one or two years of Spanish proficiency together with English proficiency, English and Spanish language arts, English and Spanish mathematics results, and Spanish social studies and science data will allow the district to see outcomes for a true 50/50 model.

CAUTIONS ON USE

CAUTION	WHAT THE TEAM STATES
CAUSAL LANGUAGE	No design in this evaluation supports causal attribution to the <u>program model</u> . There is no randomization, no matched comparison program, and, because <u>WIDA ACCESS for ELLs</u> is administered only to <u>English learners</u> , no within-district contrast on English proficiency at all. The recommendation is a judgment informed by descriptive evidence. Findings are based on converging descriptive evidence and do not establish cause.
USE AND NON-USE	This model recommendation does not reflect on the staff who implemented the current model. The allocation gap accumulated structurally, block by block over time. English outcomes under this program are strong by any national standard. This chapter recommends changing a schedule and exploring or adapting instruments. It does not recommend staffing changes and should not be cited as support for them.
EVALUABILITY	The evaluators could not compare the program's outcomes for a 50/50 versus 80/20 <u>language allocation</u> model because the program as it is currently does not reach the 50/50 allocation.

VII. PRIORITIZED RECOMMENDATIONS FOR PROGRAM IMPLEMENTATION

This section sets out what the evaluation evidence means for putting the recommendations into practice, states the strengths and conditions that the program can build on, and identifies the short-term, medium-term, and long-term action steps for implementation.

★ STRENGTHS TO BUILD ON

The evaluation found strengths that give the program a solid base for the work ahead.

- **Strong English language proficiency growth.** English learners gained an average of 76 ACCESS scale score points over two years. Three of four grade cohorts met or exceeded the two-year consortium benchmark.
- **Strong academic growth in English.** Students' MAP growth in English reading and mathematics was above the national norm in all three years. Spring 2025 IAR results exceeded the state average in every grade and subject.
- **High-quality materials in Spanish.** Science, Social Studies, Spanish language arts and mathematics Number Corner curricular materials were rated favorably for quality of content and for authentic Spanish-language.
- **Equal time by design for literacy.** The published schedule gives English and Spanish equal minutes for foundational skills, reading and writing at every grade.
- **Experienced teachers with established practice.** The dual language teachers observed have built their instructional practice over many years. That experience is the foundation for a shared, program-wide approach.
- **A linguistically rich student body.** Program-wide, partner-language composition falls within the two-way band. Students represent a heterogeneous group of language learners.
- **Data and tools already available.** The district has three years of continuous English language proficiency and academic achievement data (math and language arts). It already licenses NWEA Mathematics in Spanish and has already begun to collect and interpret Spanish and English writing samples.

CONDITIONS FOR CONTINUED SUCCESS

This section sets the conditions necessary for building on the identified strengths of the dual language program.

Language balance

Multiple data sources indicate that the language allocation for Spanish is below the threshold for a 50/50 dual language program model. The program can continue to develop and grow its bilingual learner portrait and use this as baseline quantitative and qualitative data for documenting growth.

Enrollment planning

Plans for 2027–28 should use current enrollment projections instead of the 2025 roster (V.5-F9). Dual language enrollment fell from 695 in 2020–21 to 523 in 2024–25, and it is projected at 489 and 447 for the next two years. A plan built on the 492 students in the 2025 roster would overstate the program by about 10 percent, or roughly three class sections.

Instruction

Dual language teachers have developed a cohesive set of bilingual practices, and are continuing to build skills and knowledge through effective ongoing professional learning and coaching. Pacing, materials coordination, and cross-linguistic practice, however, currently vary across classrooms and do not yet follow one coordinated program design.

Measuring biliteracy

The district has assessment data to measure Dual language students' English proficiency and English language arts to then add on Spanish language proficiency and academic achievement in Spanish.

RECOMMENDATIONS

SHORT-TERM ACTIONS · 2027-2028

The 2027-2028 school year offers an early opportunity to implement the recommendations presented in the Deliverables Section VI of this report. The district could choose one grade level or grade level band in which to adjust a schedule, pilot an assessment, or try a new rotation without affecting the whole program. There are aspects of the team's recommendations that have already started and can continue. Many recommendations can begin in the 2027-2028 school year. The following actions could be completed, or begun in the 2027-2028 school year:

1. **The schedule.** Re-adjusting and approving any schedule changes during this year would allow a revised master schedule to be in place for 2028-2029 (VI.4-R11). A
2. **The program model position.** The Board may wish to reaffirm the the 50/50 model as the district's official program model position.(VI.4-R1).
3. **Assessments in Spanish.** The district could decide by December 2027 to use MAP Lectura for native Spanish speakers only (VI.4-R6). Formally adopting the WIDA Spanish language development standards (Marco DALE) alongside the English standards (VI.3 R 2.1) would provide the basis for naming a K-5 Spanish language proficiency measure of record covering all four language domains (listening, speaking, reading and writing) (VI.4-R5), with a first administration in spring 2028. Grade-level teams could also design and pilot common Spanish oral language and writing samples, with protected time to interpret them together (VI.3-R2); choosing one grade level or grade level band as the pilot setting. A usable Spanish trend would then be available beginning in 2029.
4. **Intake screening and the linguistic portrait.** The district may wish to continue to refine the language screening and placement process and, (VI.1-R6). A single, continuing process or measure to capture students' language use in both languages could be specified by spring 2028 in coordination with the biliteracy assessment work (VI.1-R5), with linguistic portrait data available by March 2028to inform grade-by-grade rosters. (VI.4-R10).
5. **Class size and sectioning guidelines.** Before class sections are built for the 2027-2028 school year, the district could consider a planning target of twenty students per section, allocated according to student need (VI.1-R1); at least two sections at every grade (VI.1-R2); the one-third partner-language floor applied grade by grade (VI.1-R4); and a ceiling of 24 students with an agreed response when a section reaches it, in place before 2027-28 registration (VI.1-R9). Class sections should be planned on current enrollment projections rather than the 2025 roster (V.5-F7). The parent request for a class-size cap could be answered through a family-facing explanation of composition targets in English and Spanish that would help families understand the guidelines before they are adopted.

6. **Curriculum teams.** K–5 grade-level teams could continue the English–Spanish crosswalk work begun by Teaching & Learning (VI.2-R5) and begin coordinating materials and translanguaging components (across all dimensions of academic language)(VI.2-R2). Keeping foundational skills instruction within the 20 to 40 minutes recommended in the published scope and sequence would return time to the Spanish science and social studies blocks. Follow dual language program foundational skills scope & sequence (VI.2,R 4 and 5).

MIDDLE-TERM ACTIONS · Beginning 2028-2029

The 2028-2029 school year would be the first under the revised schedule. Its focus is on delivering the stated 50/50 model and gathering the first Spanish evidence. The work builds on strengths documented in the evaluation, including a literacy core that already delivers an equal 50/50 split, high-quality Spanish materials, and experienced teachers whose practice can become the program’s shared approach.

1. **Implementing and monitoring the revised schedule.** The reassigned content and specials blocks could take effect in August 2028, with more opportunities for Spanish in non-instructional settings (VI.4-R2). The allocation leads could track designed against delivered allocation through a brief schedule review or observation cycle (VI.2-R11), with a first delivered-allocation figure by grade and by block reported in June 2029 (VI.4-R3).
2. **Building sections with the adopted guidelines.** Sections for 2028-2029 could follow the 20-student target, the two-section floor, the grade-level one-third floor, and the 24-student ceiling (VI.1-R1, R2, R4, R9), with the decision rule applied if the applicant pool shifts in either direction. Information from the bilingual language use and bilingual portraits measure could be collected and used for the 2028-2029 intake. Proficiency-balanced grouping, currently a teacher-designed practice, could become a documented, program-wide practice supported by instructional coaches (VI.1-R7). Setting a minimum group size for publishing grade-level composition figures before the first report would protect small groups of students.
3. **Coordinating the two curricula.** The district may wish to adopt a version of the dual language biliteracy curriculum framework (VI.2-R7). Teacher teams could complete the crosswalk across science, social studies, and mathematics, including Number Corner, so that complementary content is taught for transfer (VI.2-R5). They could build cross-linguistic connections for all dimensions of academic language in every content area (VI.2-R2), include students’ and families’ funds of knowledge and cross-cultural content, incorporate sociocultural contexts (VI.2-R3), and add daily oral language opportunities in both languages (VI.2-R6). Coordinate materials for the content area blocks to support the move toward a true 50% English/ 50 % Spanish language allocation(VI.4-R8).
4. **Professional learning.** A translanguaging and cross-linguistic professional learning refresh could be rolled out (VI.2-R10) in coordination with the teacher team/ Teaching & Learning work on embedding crosslinguistic connections to the curricula across content areas. The coordinated writing work in grades 4 and 5 has already started in K-3 and should continue (VI.2-R8), and peer coaching would allow the program’s practitioners to share their practices with one another (VI.2-R9).
5. **Building the biliteracy assessment system from the pilot.** The district could adopt an assessment framework organized by purpose and language, with a year-long calendar reviewed by a committee of teachers, coaches, and administrators (VI.3-R1). The piloted Spanish oral language and writing protocols could be extended across grades, together with K–2 foundational literacy benchmarks and a cross-linguistic transfer inventory, and scored with a uniform rubric, calibrated team interpretation, and double-scoring of a subset of samples (VI.3). Recording and automatic transcription tools could support this work (VI.3). The district may also wish to decide on an interim language proficiency

approach (VI.3), review annual Spanish language arts tests, including a K-2 screening (VI.3), pilot a common-prompt procedure that assesses Spanish and English language arts together (VI.3), begin co-constructing biliteracy portraits on a trial basis (VI.3-R5), and expand the central data archive to hold Spanish and English evidence side by side (VI.3)).

6. **First Spanish language proficiency results alongside English results.** An annual Spanish language proficiency measure of record could be administered for the first time in the 2028-2029 school year (VI.4-R5). In the same spring, the ACCESS matched-cohort and MAP conditional-growth results could be reviewed as the English baseline (VI.4-R4), and the first annual survey on the assessment framework collected from administrators, teachers, families, and students (VI.3-R6).

LONG-TERM ACTIONS · 2029-2030 AND BEYOND

From 2030 onward, the focus shifts to sustaining the program by keeping the annual review cycles in place and strengthening the Spanish evidence base so that decisions about the program rest on a full picture of student learning in both languages.

1. **Annual review cycles.** Each year, the district could report the delivered-allocation figure by grade and block (VI.4-R3; VI.2-R11), review the English baseline each spring (VI.4-R4), build sections using the grade-level floor and decision rule (VI.1-R4), revisit how students are assigned to speaker groups (VI.1, Cautions on use), and survey administrators, teachers, families, and students on the assessment framework (VI.3-R6).
2. **A Spanish trend line.** A second administration of the Spanish proficiency measure in spring 2030 would give the program its first Spanish trend. As results accumulate, Spanish and cross-linguistic growth could be reported separately for primary English speakers, heritage Spanish speakers, and primary Spanish speakers (VI.3-R3).
3. **Mathematics in Spanish.** NWEA Mathematics in Spanish could be administered once the program reaches a true 50/50 allocation (VI.4-R7). If more mathematics instruction takes place in Spanish, the district may wish to adopt a Spanish achievement measure that reports on the same metrics as the IAR (Appendix D)
4. **Classroom-based assessment.** Once common writing samples (K-5) and other classroom-based assessments are interpreted and scored with an agreed upon process and criteria, the district could consider replacing interim NWEA administrations with monthly samples embedded in the Spanish curricular modules. Biliteracy portraits could be stored digitally, with students choosing some of their own entries (VI.3-R4 and 5).
5. **Instructional coherence.** Coherence could be sustained through teacher self-assessment, peer coaching, administrator walkthroughs, coaching cycles, and annual reviews of materials and implementation fidelity (VI.2, implementation recommendations), with professional learning continuing across the years of implementation.

VIII. LIMITATIONS OF THE EVALUATION STUDY

This evaluation was conducted between March and August 2026, using three years of district assessment data, a single week of classroom fieldwork, six administrator interviews, thirteen teacher focus-group records, six parent and six student focus-group sessions, a district-completed assessment inventory, and the district's own schedule, pacing and curriculum documentation. The limitations below affect how the findings and recommendations should be used.

This is a descriptive evaluation. It has no randomization, no matched comparison program, and no design element that supports attributing any outcome to the program model. Where the report says a result is consistent with something, it is interpreting several descriptive sources that point the same way. Every recommendation in Section VI is a professional judgment informed by that evidence, and each one includes the findings it builds on so a reader can test the judgment against them.

Absence of a Measure of Spanish Proficiency

This is the most important limitation of the study and limits how firmly the evaluation can draw conclusions.

There is no annual or interim Spanish language proficiency instrument of record at any grade; measure of Spanish oral language development or academic achievement in Spanish outside of language arts , and the program does not capture what a student can do across the two languages.

This has four consequences, each of which limits what the report can claim.

- **The Spanish outcome evidence is based on a single instrument.** Everything the report says about Spanish development in Section V comes from MAP Reading in Spanish. This conclusion is based on one measure that is normed on native Spanish speakers and is a measure of achievement (not proficiency) in Spanish reading. A decision on replacing it has been on hold since early 2025 pending this evaluation.
- **The Spanish comparison is to a self-selected population.** MAP Reading in Spanish is referenced to user norms drawn from the districts that choose to give it, largely other dual language and bilingual programs, not to a nationally representative sample. Spanish and English percentiles are therefore not on the same scale.
- **The Spanish writing evidence is based on seven samples:** four Spanish and three English pieces, at grades 4 and 5, interpreted by a single rater. The writing findings at V.2 are stated on that basis and are provisional. They are reported because they are consistent with what teachers and parents said.
- **One-third of designed mathematics time if assessed is not reported or discussed.** Mathematics is taught in two languages and tested in one (V.3-F9). Whatever is learned in the Spanish mathematics block has not been formally captured, so the report cannot say whether it is working.

A single field week

Classroom observation was conducted over one week in April 2026. The observations captured a limited sample of lessons, and they did not cover all subjects and settings evenly.

Mathematics was observed in one of the seventeen blocks, science in one, and social studies in none. All three intervention blocks were observed and all five specials ran in English. The curriculum findings therefore rest on schedule analysis, the materials review and the interview corpus. The observed allocation figure of roughly seventy-thirty is a snapshot from one week in April, and it is reported alongside three other allocation measures because no single measure is sufficient by itself and therefore reported in the aggregate. Composition in practice, who actually speaks which language to whom, was not captured. A single week in April also has a seasonal character that the report does not attempt to correct for. April falls after state testing at some grades and inside it at others, and a schedule in April is not necessarily the schedule in October.

Evidence and professional judgment

- **The recommendations address program conditions.** Nothing in this report supports a conclusion about the competence or commitment of the staff who have run this program. The allocation shifts have accumulated structurally over time, and the teachers delivered a faithful 50/50 literacy core, which was the part of the model they were responsible for. English outcomes under this program are strong by any national standard.
- **Most of these limitations can be addressed.** The need for measuring Spanish academic language in a systematic manner, the unrecorded language of administration, the missing curriculum documentation, the need for a more refined initial language proficiency protocol, and the unconfirmed room count are all conditions the district can change, and each is addressed by a numbered recommendation in Section VI.

GLOSSARY

NORTH SHORE SCHOOL DISTRICT 112 · DUAL LANGUAGE PROGRAM EVALUATION

Glossary · Dual Language and Assessment Terms

Paridad Education Consulting ·

HOW TO USE THIS GLOSSARY

This glossary defines the terms of bilingual education and educational measurement that recur throughout the evaluation report. The first section defines the concepts. The second defines the tests the report names and the types of scores each uses.,

TERMS

TERM	WHAT IT MEANS
A	
Academic achievement	What a student knows and can do in a subject — reading, mathematics, science. Distinct from language proficiency. A test given in English measures both at once, which is why results in a dual language program are read alongside the language of the test.
B	
Bilingualism	The ability to understand and use two languages. In a dual language program it is a goal for every student, not only for students who arrive speaking the partner language.
Biliteracy	The ability to read and write in two languages. Bilingualism covers understanding and speaking; biliteracy adds reading and writing, and it develops on a longer timeline. It is one of the three pillars of dual language education.
Cross Linguistic Transfer	Planned instruction that connects what students learned in one language to the other; for example, naming in English the concepts and vocabulary of a unit taught in Spanish. Bridging is planned as part of the lesson and is different from incidental translation.
C	
Cohort	A group of students followed together over time, such as the students who were in grade 3 in 2022–23 and grade 5 in 2024–25. Reporting by cohort shows whether the same children grew, rather than comparing one year's third graders to the next year's.

TERM	WHAT IT MEANS
Composite score	A single score that combines several parts of a test, weighted by the test maker. An English proficiency composite blends listening, speaking, reading, and writing. Because it is a blend, a rising composite can conceal a flat score in one domain.
Conditional growth percentile	How much a student grew compared with students nationally who started the year at the same score. Fifty is average growth. It answers “did this child grow more than similar children,” which is a different question from “how high did this child score.”
Cross-sectional	A snapshot of whoever is enrolled at one point in time — for example, all of this year's third graders. Useful for describing a moment; it cannot show whether individual students grew. Contrast with longitudinal.
Curriculum coherence	Whether what is taught in each language, in each grade, fits together as one program instead of as two separate programs, in two different languages in the same building.
D	
Disaggregation	Breaking results apart by group — grade, school, program, language proficiency — instead of reporting a single overall average. Averages can hide a pattern that shows up clearly once the groups are separated.
Language Preference	The language in which a student currently performs most strongly. is not the same as the language spoken at home, and it is not the same as English learner status. Establishing it requires assessing both languages.
E	
Effect size	A number expressing how large a difference is, independent of whether it is statistically detectable. Effect sizes let findings from different studies be compared, but they require measures built for that purpose.
English learner (EL)	A state classification for a student whose English proficiency is still developing, assigned on the basis of a home language survey and a proficiency screener. It is a label about English proficiency, not about a student's home language, ability, or program placement, and the label is removed upon reclassification.
F	
Fidelity of implementation	The degree to which a program is delivered as it was designed. A model can be sound on paper and still produce weak results if what reaches classrooms differs from the plan.
G	
Guiding Principles for Dual Language Education	The field's standard framework for designing and evaluating dual language programs, published by the Center for Applied Linguistics and organized into strands of practice .
H	

TERM	WHAT IT MEANS
Heritage speaker	A student who grew up hearing or speaking a language at home, often with strong oral command of it and less developed academic reading and writing in that same language.
I	
Interim assessment	A test given at set points in the year — commonly fall and spring — to track progress between the daily checks of formative assessment and the year-end judgment of a summative test. Also called a benchmark assessment.
L	
Language allocation	How instructional time is divided between the two languages: which subjects are taught in which language, for how many minutes, at each grade. Published allocation is what the plan states; delivered allocation is what the master schedule actually produces, and the two can differ.
Language arts	The school subject that covers achievement in reading, writing, speaking, and listening. In a dual language program there are two such subjects: English language arts (ELA) and Spanish language arts (SLA) that are parallel subjects with their own standards and their own instructional time.
Language development	The process by which a student's command of a language grows over time across all four domains (listening, speaking, reading and writing). In dual language research it is understood to advance alongside academic achievement, thinking, and cultural understanding rather than ahead of them. Students do not need to master one language before learning content in it.
Language domains	Listening, speaking, reading, and writing. Proficiency is measured domain by domain. Listening and speaking together are usually reported as oral language; reading and writing together as literacy.
Language proficiency	How well a student can use a language, both for conversation and for school tasks, reported by domain. Distinct from academic achievement, which measures content knowledge.
Longitudinal	Following the same students across years, so that change within a child can be observed. Contrast with cross-sectional.
LOTE	Language Other Than English in a dual language program
M	
Multilingual learners	The broader term for students developing more than one language, whether or not they carry the English learner label. It is increasingly preferred because it names what a student has rather than what a student lacks, and because it includes children who are bilingual from birth and were never classified as English learners.
N	

TERM	WHAT IT MEANS
Norm-referenced and criterion-referenced assessment	Two ways of giving a score meaning. Norm-referenced compares a student with other students (percentiles). Criterion-referenced compares a student with a fixed standard of what students should know (proficiency cut scores). Each answers a different question.
Norms	The reference group a score is compared against. National norms compare a student with a nationally representative sample. User norms compare a student with everyone who happens to take that test, which is not a representative group. Scores built on different norms cannot be subtracted from one another.
O	
P	
Pacing calendar	The grade-by-grade schedule setting out which unit is taught when. In a dual language program it also determines whether the two languages are taught in the same week or alternate, and whether grades move through the sequence together.
Partner language	The language of instruction other than English:Spanish, in SD 112. The term indicates that Spanish has equal standing with English.
Portfolio and work samples	Collected student work — writing samples, projects, recorded speech; interpreted and scored against a rubric. Portfolios can show what a multiple-choice test cannot, particularly in writing and in the partner language.
Proficiency level	A labeled band describing what a student can typically do in a language, such as levels 1 through 6. The levels are ordered, but the distance between them is not equal and they are not grades.
Proficiency level descriptors	The written statements of what a student at each language proficiency level can typically do, in each domain (listening, speaking, reading and writing). They give each level a specific meaning.
Program model	The overall design of a program: which languages, which subjects in each, how much time, which grades. 50/50 and 80/20 are program models. Changing the model is a different decision from changing the curriculum.
R	
Reclassification	The point at which a student's English proficiency score meets the state criterion and the English learner label is removed. In a dual language program reclassification changes only the label. The student keeps the same seat, the same section, and the same two languages of instruction. Because the label disappears as students succeed, it should not be used as a measure of who speaks what.

TERM	WHAT IT MEANS
RIT score	The scale on which the reading and mathematics achievement tests used here report results. A RIT score carries the same meaning in any grade, so an individual student's scores can be followed across years. It is not a percentage, not a percentile, and not a grade level. This report does not compare RIT scores across years directly; it uses conditional growth percentiles for that.
Rubric	A scoring guide describing what performance looks like at each level. Rubric scores are locally interpreted; they provide useful information about a student's work, but they are not on a common scale with test scores and should not be combined with them.
S	
Scale score	A score on a continuous scale that holds the same meaning across grades and years, so growth can be tracked over time. The reading and mathematics scale used in this report is called RIT. Scale scores are not percentages and not percentiles.
Screener	A short assessment given to all students to identify who may need further attention, or, in the case of language, to determine an initial classification. A screener identifies students for follow-up but does not diagnose.
Seal of Biliteracy	A state-awarded recognition on a high school diploma for students who demonstrate proficiency in English and at least one other language. In Illinois it is the formal endpoint of the K–12 biliteracy pathway a dual language program begins.
Sociocultural competence	The third pillar of dual language education: students' ability to move between cultures with understanding and respect, and to hold a positive sense of identity in both languages. It is the pillar least often measured, because it calls for observation and student voice rather than a test score.
Standards	Academic content standards and language development standards
Statistical significance	An indication that a difference is unlikely to be the product of chance. It says nothing about whether the difference is large enough to matter in a classroom. That question concerns practical significance, which is a separate judgment.
Strands of Practice	In the Guiding Principles framework, one of the areas of practice a program is examined against — curriculum, instruction, assessment and accountability, staff quality and professional development, program structure, support and resources, and family and community. Findings in this report are mapped to strands.
Student Growth Percentiles	
Summative assessment	An end-of-period test reporting what students learned. State accountability tests are summative. They are designed for reporting and are less useful for adjusting day-to-day instruction.
T	

TERM	WHAT IT MEANS
Three pillars of dual language education	High academic achievement, bilingualism and biliteracy, and sociocultural competence. A dual language program is judged against all three together; strong test scores alone do not make a program successful under this framework.
Transitional bilingual education (TBE)	A model that uses a student's home language temporarily as a bridge into all-English instruction. Its goal is English proficiency, and the home language is phased out. In contrast, a dual language program develops and maintains both languages.
Translanguaging	A student's natural use of their full linguistic range, drawing on both languages to make meaning. As a pedagogy it means planning for that deliberately, including when each language is used on its own, instead of treating a student's mixing of languages as an error.
Two-way immersion	A dual language program enrolling both students who speak the partner language at home and students who speak English at home, taught together in the same classrooms. The field commonly looks for at least about one third of a class from each group, so that students serve as language models for one another. SD 112's program is two-way.
V	
Vertical articulation	Whether what is taught in one grade builds on the grade below and prepares for the grade above — in both languages. Weak vertical articulation shows up as gaps and repetition that are hard to see from within a single grade.

ASSESSMENTS NAMED IN THIS REPORT

Each entry says what the assessment measures, in which language, and what its results can and cannot support.

TERM	WHAT IT MEANS
WIDA ACCESS for ELLs	The annual English language proficiency test Illinois requires for every student identified as an English learner, kindergarten through grade 12. It reports listening, speaking, reading, and writing, and combines them into oral language, literacy, and an overall composite score on a 1 through 6 scale. Illinois sets the reclassification criterion at a composite of 4.8. This evaluation analyzed three years of it, 2022–23 through 2024–25.
FastBridge	Early literacy and oral reading measures given in English at kindergarten through grade 2, fall and spring, used here to describe early literacy trajectory. Spanish forms exist in the same product line
Illinois Assessment of Readiness (IAR)	The state accountability test in English language arts and mathematics, grades 3 through 5, given each spring in English. It reports the share of students at each proficiency level against Illinois cut scores, together with student growth percentiles.

TERM	WHAT IT MEANS
Marco DALE	A WIDA framework for describing Spanish language proficiency: a six-level continuum applied across three dimensions — discurso, oración, and palabra/frase. It describes where a student falls on a Spanish development continuum, which is a different question from whether a piece of work meets a grade-level standard. The district does not use it. The evaluation team applied it to a small set of Spanish writing samples because the district has no Spanish proficiency measure of record.
NWEA MAP Lectura	The Spanish-language counterpart to MAP, and the district's only Spanish-language achievement measure. It is reported on user norms rather than national norms, which supports statements about trend and growth but rules out subtracting a Spanish score from an English one. This report makes no such comparison.
NWEA MAP — Reading and Mathematics	Adaptive achievement tests given in English, fall and spring. This report gives priority to conditional growth percentile over achievement percentile.
WIDA SCREENER	The multistate consortium whose English language development standards and assessments Illinois uses. Both ACCESS for ELLs and the Marco DALE descriptors are WIDA frameworks. When this report benchmarks growth against “WIDA Consortium averages,” it means the average across all states using the assessment.

REFERENCES

- Bibler, A. (2021). Dual language education and student achievement. *Education Finance and Policy*, 634–657. https://doi.org/10.1162/edfp_a_00320
- Collier, V. P., & Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale, longitudinal research. *Annual Review of Applied Linguistics*, 37, 203–217.
- English Learners Success Forum. (n.d.). Guidelines for improving English language arts materials for English learners.
- Fortune, T. W. (2015). Oral proficiency assessment of English-proficient K–8 Spanish immersion students. *The Modern Language Journal*, 99(4), 637–655.
- Fortune, T. W., & Tedick, D. J. (Eds.). (2008). *Pathways to multilingualism: Evolving perspectives on immersion education*. Multilingual Matters.
- Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (Eds.). (2006). *Educating English language learners: A synthesis of research evidence*. Cambridge University Press.
- Heineke, A. J., & Davin, K. J. (2025). Pathways to biliteracy: Investigating the implementation of the Seal of Biliteracy in elementary and secondary schools in Chicago. *Urban Education*, 60(6), 1672–1707.
- Heath, C., & Heath, D. (2010). *Switch: How to change things when change is hard*. Crown Business.
- Hilliard, J. F., & Gottlieb, M. (2021, June) Seeing multilingual learners' reflections en EL ESPEJO. *Bilingual Basics*, the newsletter of the Bilingual Multilingual Education Interest Section of TESOL International Association.
- Howard, E. R., Lindholm-Leary, K. J., 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.
- Howard, E. R., & Sugarman, J. (2009). Program models and the language of initial literacy in two-way immersion programs. Center for Applied Linguistics.
- Howard, E. R., Sugarman, J., & Christian, D. (2003). Trends in two-way immersion education: A review of the research (CREDE Report No. 63). Center for Research on Education, Diversity & Excellence.
- Lindholm-Leary, K. J. (2004). The rich promise of two-way immersion. *Educational Leadership*, 62(4), 56–59.
- Lindholm-Leary, K. (2012a, April). Background and achievement of English language learners with different levels of English proficiency in dual language programs [Conference paper]. American Educational Research Association Annual Meeting.
- Lindholm-Leary, K. (2012b). Success and challenges in dual language education. *Theory Into Practice*, 51(4), 256–262.
- Lindholm-Leary, K. (2019). English language proficiency trajectories of Latino EL students in dual language programs. *NABE Journal of Research and Practice*, 9(2), 78–87.
- Lindholm-Leary, K., & Borsato, G. (2005). Hispanic high schoolers and mathematics: Follow-up of students who had participated in two-way bilingual elementary programs. *Bilingual Research Journal*, 29(3), 641–652.
- Mihaly, K., Arellano, B., & Prier, S. (2022). Biliteracy seals in a large urban district in New Mexico: Who earns them and how do they impact college outcomes? (REL 2023–140). Regional Educational Laboratory Southwest, Institute of Education Sciences.
- National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. The National Academies Press. <https://doi.org/10.17226/24677>
- Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review*, 84(1), 85–100.

RAND Corporation. (2017). Dual-language immersion programs raise student achievement in English (Research Brief RB-9903). RAND Corporation.

Sánchez López, C., Young, T., Genesee, F., & Hilliard, J. F. (2022). Welcoming bilingual learners with disabilities into dual language programs: Research-informed implications for inclusive programs and practice. Center for Applied Linguistics / National Dual Language Forum / Paridad Education Consulting.

Slavin, R. E. (1989). Class size and student achievement: Small effects of small classes. *Educational Psychologist*, 24(1), 99–110.

Steele, J. L., Slater, R. O., Li, J., Zamarro, G., & Miller, T. (2015). Costs and effects of dual-language immersion in the Portland Public Schools [Conference paper]. Society for Research on Educational Effectiveness Spring Conference. RAND Corporation.

Steele, J. L., Slater, R. O., Li, J., Zamarro, G., & Miller, T. (2018). Dual-language immersion education at scale: Analysis of program fidelity, achievement, and mediating factors. *Educational Evaluation and Policy Analysis*, 40(3), 420–445.

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_suppl), 282S–306S.

Thomas, W. P., & Collier, V. P. (2002). A national study of school effectiveness for language minority students' long-term academic achievement. Center for Research on Education, Diversity & Excellence.

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.

Valentino, R. A., & Reardon, S. F. (2015). Effectiveness of four instructional programs designed to serve English learners: Variation by ethnicity and initial English proficiency. *Educational Evaluation and Policy Analysis*, 37(4), 612–637.

APPENDICES

Appendix A Evaluation Question Maps/ Guiding Principles Crosswalk

EQ 1 Map-English Language Proficiency (ELP)

EQ 1: To what extent are students developing English language proficiency, and are English Learners progressing toward reclassification? (RFP Sec. C)

In operational terms: How much English language proficiency did English Learners gain each year from spring 2023 to spring 2025, measured on the WIDA ACCESS for ELLs composite scale score, and how many students were reclassified out of English Learner status over the same span against the Illinois 4.8 criterion?

Primary Focus: ACCESS for ELLs™ data and reclassification rates for ELs.

Guiding Principles		Three Pillars	
SEVEN STRANDS (the means) advance → the pillars (the goals)	Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	English exposure set by the allocation model	Articulated PreK–12 grade-level pathway	Equal status for both languages by design
2 Curriculum	English Language Development (ELD) & English language arts (ELA)	Standards-aligned content in both languages	Culturally sustaining, multicultural content
3 Instruction	Practices that build English language proficiency	Scaffolded, cognitively rigorous content	Affirms students' languages & identities
4 Assessment & Accountability	ACCESS data & reclassification rates; Enrollment & demographic data	Monitors grade-level progress; guides instruction	Tracks cross-cultural growth; equitable measures
5 Staff Quality & Prof. Development	ELD professional learning; biliterate teachers	Two-language content & pedagogy expertise	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures
7 Support & Resources	Time & materials for the partner language	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 1 ↓

DELIVERABLE

Instructional Model Analysis & Recommendations

DATA SOURCES · EQ 1

QUANTITATIVE

- ACCESS for ELLs (WIDA)
- Reclassification records (4.8 criterion)
- Enrollment & demographic data

QUALITATIVE

- Classroom observations (rubric)
- Teacher interviews & focus groups
- Administrator interviews
- Document review (reclass. policy)
- Professional development plan

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

EQ 2 Map- Spanish Language Proficiency (SLP)

EQ 2: To what extent are all students developing Spanish language proficiency? (RFP Sec. C, D)

In operational terms: *What do grade 4–5 Spanish writing samples, classroom observations, curriculum review, and teacher, parent, and student focus groups show about students' Spanish proficiency and about the conditions producing it?*

Primary Focus: Assessment of Spanish language proficiency and literacy

Guiding Principles		Three Pillars	
SEVEN STRANDS (the means) advance → the pillars (the goals)			
	Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Spanish time protected by the allocation model	Articulated PreK–12 grade-level pathway	Equal status for both languages by design
2 Curriculum	Integrated Spanish Language Development (SLD) & Spanish Language Arts (SLA)	Standards-aligned content in both languages	Culturally sustaining, multicultural content
3 Instruction	Practices to optimize Spanish language proficiency	Scaffolded, cognitively rigorous content	Affirms students’ languages & identities
4 Assessment & Accountability	Assessment of Spanish-language proficiency & literacy	Monitors grade-level progress; guides instruction	Tracks cross-cultural growth; equitable measures
5 Staff Quality & Prof. Development	Biliterate teachers fluent in Spanish	Two-language content & pedagogy expertise	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Access to community knowledge; bridges cultures
7 Support & Resources	Time & materials for Spanish language	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 2 ↓

DELIVERABLES

Recommendations for a Comprehensive Biliteracy Assessment Framework

DATA SOURCES · EQ 2

QUANTITATIVE

- Spanish LP measure
- Spanish writing-sample ratings (gr 4–5)
- Enrollment & demographic data

QUALITATIVE

- Language-sample review (gr 4–5)
- Classroom observations (Spanish)
- Teacher interviews & focus groups
- Professional Development Plan
- Materials review

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

EQ 3 Map - Academic Achievement

EQ 3: What is the achievement of EL students in the Dual Language program for Language Arts and Mathematics? (RFP Sec. C)

In operational terms: *What do the data show about achievement in the dual language program on the annual state test, and the interim measures for mathematics and language arts separately, English and Spanish separately, across the three years for which data exist?*

Primary Focus: State and district Language Arts and Math annual and interim standardized tests

Guiding Principles		Three Pillars	
SEVEN STRANDS (the means) advance → the pillars (the goals)	<u>Bilingualism & Biliteracy</u>	<u>Academic Achievement</u>	<u>Sociocultural Competence</u>
1 Program Structure	<u>Language allocation</u> protects partner-language time	Grade-level content across both languages	Equal status for both languages by design
2 Curriculum	<u>Biliteracy</u> trajectory; cross-linguistic bridge	Standards-based content in both languages	Culturally sustaining, multicultural content
3 Instruction	Strategies for transfer & target-language output	High quality content instruction in both languages	Affirms students' languages & identities
4 Assessment & Accountability	Assesses proficiency & literacy in both languages	State and district <u>Language Arts/Math</u> tests	Tracks cross-cultural growth; equitable measures
5 Staff Quality & Prof. Development	Biliterate, dual-language trained teachers	Content + pedagogy expertise (English & Spanish)	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures
7 Support & Resources	Time & materials for the <u>partner language</u>	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 3 ↓

DELIVERABLE

Instructional Model Analysis & Recommendations

DATA SOURCES · EQ 3

QUANTITATIVE

- IAR — ELA & Math (grades 3–8)
- MAP Growth — English & Spanish
- Enrollment & demographic data
- content area assessment data in Spanish

QUALITATIVE

- Classroom observations (content)
- Teacher focus groups
- Instructional materials review
- Materials Review
- Document review (curriculum maps)

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

EQ 4 Map - Biliteracy Assessment Framework

EQ 4: How might a comprehensive biliteracy assessment framework be conceptualized for the dual language program in SD 112? (RFP Sec. C, D)

In operational terms: *What data and what instrumentation are in place to capture dual language students' literacy and oral language (listening and speaking) development in English, in Spanish, and across both languages?*

Primary Focus: *How might SD 112 plan for a biliteracy trajectory with crosslinguistic connections to promote language and literacy development?*

Guiding Principles		Three Pillars	
SEVEN STRANDS (the means) advance → the pillars (the goals)	Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Language allocation	Allocation gives content time in Spanish	Equal status for both languages by design
2 Curriculum	Evidence of <u>biliteracy</u> development with cross-linguistic connections	Standards-aligned content in both languages	Culturally sustaining, multicultural content
3 Instruction	Evidence of <u>biliteracy</u> strategies with cross-linguistic connections	Grade level Content instruction taught in Spanish	Affirms students' languages & identities
4 Assessment & Accountability	Assessment of proficiency & literacy in both languages	Standardized and <u>standards</u> referenced Spanish content assessment data	Tracks cross-cultural growth; equitable measures
5 Staff Quality & Prof. Development	Biliterate, dual-language trained teachers	Teachers deliver content in academic Spanish	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures
7 Support & Resources	Time & materials for the <u>partner language</u>	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 4 ↓

DELIVERABLES

Recommendations for a Comprehensive Biliteracy Assessment Framework

also feeds: Instructional Model Analysis (EQ 1–6)

DATA SOURCES · EQ 4

QUANTITATIVE

- MAP Lectura (Spanish reading)
- Spanish NWEA Mathematics
- English language development measures (WIDA ACCESS)

QUALITATIVE

- Classroom observations (Spanish)
- Teacher, Family and Student focus groups
- Instructional materials review (Spanish and English)
- Spanish & English writing samples Grades 4/5

EQ 5 Map - Class Size

EQ 5: What are the optimal class sizes and student composition ratios for K–5 dual language classrooms in SD 112? (RFP Sec. B)

In operational terms: *Do the class sizes and composition ratios currently in place support the two-way model, and what configurations does the evidence indicate would strengthen it?*

Primary Focus: Program Structure, class size and speaker-composition ratios as the structural conditions that preserve the two-way model

Guiding Principles		Three Pillars	
SEVEN STRANDS (the means) advance → the pillars (the goals)	Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Student ratios to promote bilingual development	Maintenance of balance between student groups	Equal status for both languages
2 Curriculum	<u>Biliteracy</u> trajectory; cross-linguistic bridge	Standards-aligned content in both languages	Culturally sustaining, multicultural content
3 Instruction	Strategies for transfer & target-language output	Scaffolded, cognitively rigorous content	Affirms students' languages & identities
4 Assessment & Accountability	Assesses proficiency & literacy in both languages	Monitors grade-level progress; guides instruction	Tracks cross-cultural growth; equitable measures
5 Staff Quality & Prof. Development	Biliterate, dual-language trained teachers	Two-language content & pedagogy expertise	Culturally responsive, diverse educators
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures
7 Support & Resources	Time & materials for Spanish language development & Achievement	Adequate resources for dual instruction	Resources reflecting both cultures equitably

↓ EQ 5 ↓

DELIVERABLE

Recommendations for Dual Language Class Size

DATA SOURCES · EQ 5

QUANTITATIVE

- Enrollment & demographics by school/grade
- Current class sizes & composition ratios
- Oak Terrace capacity & enrollment projections

QUALITATIVE

- Administrator interviews (capacity & staffing)
- Classroom observations (composition in practice)

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

EQ 6 Map - Curriculum and Pacing Framework

EQ 6: How coherent, vertically aligned, and consistently paced is the dual language curriculum across grade levels in SD 112?

In operational terms: *Do the scope and sequence, the pacing calendars, the materials, and the language allocation add up to one coordinated K–5 program — or to two programs taught side by side?*

Primary Focus: Curriculum

Guiding Principles		Three Pillars		
SEVEN STRANDS (the means) advance → the pillars (the goals)		Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Language allocation protects partner-language time	Articulated PreK–12 grade-level pathway	Equal status for both languages by design	
2 Curriculum	Biliteracy trajectory; cross-linguistic bridge	Standards-aligned content in both languages	Culturally sustaining, multicultural content	
3 Instruction	Strategies for transfer & target-language output	Scaffolded, cognitively rigorous content	Affirms students' languages & identities	
4 Assessment & Accountability	Assesses proficiency & literacy in both languages	Monitors grade-level progress; guides instruction	Tracks cross-cultural growth; equitable measures	
5 Staff Quality & Prof. Development	Biliterate, dual-language trained teachers	Two-language content & pedagogy expertise	Culturally responsive, diverse educators	
6 Family & Community	Home-language use & literacy partnership	Engagement that supports achievement	Honors community knowledge; bridges cultures	
7 Support & Resources	Time & materials for the <u>partner language</u>	Adequate resources for dual instruction	Resources reflecting both cultures equitably	

↓ EQ 6 ↓

DELIVERABLE Curriculum & Pacing Framework Recommendation

DATA SOURCES · EQ 6

QUANTITATIVE

- Instructional minutes & language allocation by grade
- Pacing variance across sites & grade levels
- Curriculum depth and breadth

QUALITATIVE

- Materials review
- Teacher, student & family focus groups
- Classroom observations (allocation in practice)
- Administrator interviews (curriculum planning)

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

EQ 7 Map - Program Model

EQ 7: How effective is the current 50/50 dual language program model? What program model is recommended for SD 112?

In operational terms: What does the collective evidence tell us, and how does it point to the most appropriate dual language program model for English learners in District 112?

Primary Focus: To what extent are the Three Pillars operationalized across the current program structure?

Guiding Principles		Three Pillars		
STRANDS of practice (the means) advance → the pillars (the goals)		Bilingualism & Biliteracy	Academic Achievement	Sociocultural Competence
1 Program Structure	Language allocation protects 50/50 model	Articulated K–5 grade-level pathway		Equal status for both languages by design
2 Curriculum	Integrated Spanish & English <u>Language Development</u> and <u>Language Arts</u>	Standards-aligned grade level content in both languages	Linguistically & Culturally sustaining, multicultural content	
3 Instruction	Practices to optimize <u>bilingualism</u> & <u>biliteracy</u>	Cognitively engaging & demanding content in both languages	Integrates students’ linguistic and cultural resources	
4 Assessment & Accountability	Assessment of proficiency & literacy in both languages	Documentation of grade-level progress across both languages	Application of cultural competencies at the individual, classroom and school levels	
5 Staff Quality & Prof. Development	Biliterate teachers; dual language-specific professional development	Content knowledge & pedagogical expertise across both languages	Cultural & linguistic competencies interwoven into professional learning	
6 Family & Community	Home-language use & <u>biliteracy</u> development	Engagement that supports achievement	Access to community knowledge; bridges cultures	
7 Support & Resources	Time & materials for for both languages	High-quality resources for dual language instruction	Access to authentic cultural & linguistic experiences	

↓ EQ 7 ↓

DELIVERABLES

Instructional Model Analysis & Recommendations

also feeds: Curriculum & Pacing Framework (EQ 6, 8)

DATA SOURCES · EQ 7

QUANTITATIVE

- Language-allocation analysis
- Proficiency & achievement trends
- Reclassification trajectories

QUALITATIVE

- Classroom observations (allocation)
- Administrator interviews
- Teacher interviews & focus groups
- Document review (allocation plan)

Strands shown (each holds several discrete guiding principles). Framework after Howard et al., Guiding Principles, 3rd ed. (CAL, 2018).

Appendix B Data Collection Instruments

Qualitative Data Tracker

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2		SD #112 Qualitative Data Collection — Summary Dashboard												
3		Collection window: 2022–23 2023–24 2024–25 — Status counts auto-update from each tab												
4														
5		Qualitative Domain	Total Items	Collected	In Progress	Not Started	% Complete							
6		Program Design & Implementation	0	0	0	0	0%							
7		Classroom Instruction & Practice	0	0	0	0	0%							
8		School Culture & Climate	0	0	0	0	0%							
9		Family & Community Engagement	0	0	0	0	0%							
10		Leadership, Policy & Equity	0	0	0	0	0%							
11		Staffing & Professional Dev.	0	0	0	0	0%							
12		TOTAL	0	0	0	0	0%							
13		Upload all materials to the shared Google Drive folder. Store confidential materials (interviews, focus groups, etc.) in a separate folder.												
14														
15														
16														
17														
18														

Quantitative Data Tracker

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1																
2		NSSD 112 Data Collection — Summary Dashboard														
3		School years in scope: 2022–23 2023–24 2024–25 — Click a domain name to open its Google Drive folder														
4																
5		Analytical Domain	Total Items	Uploaded	In Progress	Not Started	% Complete									
6		English Language Proficiency	0	0	6	0	0%									
7		Spanish Language Proficiency	0	0	0	0	0%									
8		Achievement in English	0	0	0	0	0%									
9		Achievement in Spanish	0	0	1	0	0%									
10		Supporting / Contextual Data	0	0	0	0	0%									
11		TOTAL	0	0	0	0	0%									
12		Upload all files to the shared Google Drive folder. Enter filename(s) in 'File Name(s) in Drive' on each tab. Update the														
13																
14																
15																
16																
17																

SD 112 DL Evaluation — Instructional Materials Review Protocol

SD 112 Dual Language Program Evaluation | Paridad Education Consulting | 2025–26

Reviewer: Cristina Sánchez-López

Complete one submission per material series reviewed. Refer to the printed protocol for the full Framework Alignment section, Grade-Band Applicability Notes, and Rating Scale before beginning each review.

RATING SCALE:

3 — Strong: Clear, consistent evidence. Meets the standard without significant modification.

2 — Partial: Some evidence present but incomplete or requires teacher supplementation.

1 — Limited/Absent: Little or no evidence. Significant gap relative to the standard.

N/A — Not applicable to this material, subject area, or language context.

* Indicates required question

MATERIAL IDENTIFICATION

Complete all fields before beginning the dimension ratings.



DUAL LANGUAGE CLASSROOM OBSERVATION RUBRIC

Flow View — Criteria in Recommended Observation Sequence

North Shore School District 112 Program Evaluation

Aligned with the Guiding Principles for Dual Language Education (3rd Edition)

★ = Language Arts observation only · Extended criteria (shown after Core) are collapsed by default in the digital tool

OBSERVATION INFORMATION

Observer Name:	_____	Teacher Name:	_____
Grade Level:	_____	School:	☐ Oak Terrace ☐ Red Oak
Date:	_____	Start Time:	_____
End Time:	_____	Language of Instruction:	☐ Spanish ☐ English
Subject/Content Area:	_____	# Students Present:	_____
Native Spanish Speakers:	_____	Native English Speakers:	_____
# English Learners:	_____	EL Proficiency Levels:	_____

SD 112 Dual Language Program Evaluation
Assessment Inventory — Verification Worksheet

SD 112 DL Evaluation - Assessment Inventory Verification

Prepared for verification by Enrique Castro · District 112

Enrique — as part of the dual language program evaluation (Evaluation Question 4, the biliteracy assessment framework), we are confirming the district's current assessment landscape. Please review each instrument below and indicate whether District 112 is **currently using this assessment**. This is an inventory for verification only — no evaluative judgment is implied by an instrument's presence or absence.

Part A · District, State & Language-Proficiency Measures

Instrument (and what it measures)	Language	Grades	Frequency	Using?	Notes / to confirm
WIDA Screener Initial English language proficiency for EL identification and placement.	English	K-5	At entry	☐ Yes ☐ No ☐ Partial	
WIDA ACCESS for ELLs Annual English language proficiency across listening, speaking, reading, writing; basis for reclassification (4.8 composite).	English	K-5	Annual (winter)	☐ Yes ☐ No ☐ Partial	
NWEA MAP — Reading Interim, norm-referenced English reading growth.	English	K-5 (typ. 2-5)	2-3x / year	☐ Yes ☐ No ☐ Partial	
NWEA MAP Lectura Interim, norm-referenced Spanish	Spanish	Gr 2-5	2-3x / year	☐ Yes ☐ No	Confirm grades administered

50/50 Model Schedule Kindergarten (Weeks 1-2)	
Time	Subject Area
8:50-9:50 am 60 Minutes	Foundational Skills
9:50-10:20 am 30 Minutes	Student Success Block
10:20-11:00 am 40 Minutes	English Language Arts - Reading
11:00-11:20 20 Minutes	Spanish Language Arts - Writing
11:20-11:50 30 Minutes	Science/Social Studies
11:50-12:40 pm 50 Minutes	Lunch/Recess
12:45-1:45 pm 60 Minutes	CAPE

Appendix C Interview and Focus Group Protocols

Admin Staff Interviews .DOCX

File Edit View Insert Format Tools Help

100% Normal text Verdana 11 B I U A

Front Office Staff

Opening Script — Read Aloud

Thank you for being here. This is a 30-minute focus group as part of the SD 112 Dual Language Program Evaluation. The goal is to understand how the dual language program looks from your perspective.

This is not an evaluation of your individual work. Findings will be reported in aggregate — no individual person, classroom, or school will be identified. We ask that you treat what is said here as confidential. The Paridad Team will also keep all information confidential. Please speak candidly to your comfort level.

With your permission, I'll record this so we have an accurate transcript. The recording stays with the evaluation team. May I record?

[On verbal consent from all, start primary + backup. Then state:]

Today is [], This is [], conducting a Protocol 2 focus group at []

Voice ID Round + Team Intro — ~3 min

Engagement Specialist Interview protocol .DOCX

File Edit View Insert Format Tools Help

100% Heading 2 Verdana 11 B I U A

Engagement Specialist

Opening Script — Read Aloud

Thank you for being here. This is a 30-minute focus group as part of the SD 112 Dual Language Program Evaluation. The goal is to understand how the dual language program looks from your perspective.

This is not an evaluation of your individual work. Findings will be reported in aggregate — no individual person, classroom, or school will be identified. We ask that you treat what is said here as confidential. The Paridad Team will also keep all information confidential. Please speak candidly to your comfort level.

With your permission, I'll record this so we have an accurate transcript. The recording stays with the evaluation team. May I record?

[On verbal consent from all, start primary + backup. Then state:]

Today is []. This is [], conducting a Protocol 2 focus group at []. All

SD 112 Dual Language Program Evaluation

Parent Interview — Field Protocol

Paridad Education Consulting • Spring 2026 • Protocol 3 (English) • ~30 min • 1-on-1, Recorded

Intended for: 30-minute, 1-on-1 interviews with parents/guardians of children currently or previously enrolled in the SD 112 dual language program. Conducted in English with English-preference families. A separate Spanish-language version of this protocol exists for Spanish-preference families and should always be offered as an option — the family chooses the language.

Language note: Always confirm the family's preferred language for the interview before scheduling. The interviewer should be a fluent speaker of that language. Spanish-medium parent interviews should be conducted by Cristina or another fluent Spanish speaker on the team — never through translation by the parent's child or another family member.

Pre-Recording Checklist (Not Applicable)

- ☐ Batteries checked on primary (H1) and backup (iPad)
- ☐ SD card cleared; WAV / 44.1 kHz confirmed
- ☐ 10-second level test recorded and played back
- ☐ Backup device in airplane mode, recording started
- ☐ Printed protocol, pen, water on hand
- ☐ Participant key updated with this session
- ☐ Quiet, comfortable space confirmed (parent's preference: home, school, community location)

Evaluación del Programa de Lenguaje Dual — SD 112

Entrevista a Padres — Protocolo

Paridad Education Consulting • Primavera 2026 • Protocolo 3 (Español) • ~30 min • Individual, Grabada

Para quién es: Entrevistas individuales de 30 minutos con madres, padres o tutores de niños/as actualmente o anteriormente inscritos en el programa de lenguaje dual de SD 112. Realizadas en español con familias que prefieren ese idioma. Una versión en inglés de este protocolo existe para familias que prefieren inglés — la familia elige el idioma.

Nota sobre el idioma: Confirme siempre el idioma preferido de la familia antes de programar la entrevista. La persona entrevistadora debe ser hispanohablante con fluidez. Las entrevistas con familias hispanohablantes nunca deben realizarse a través de la traducción del hijo/a u otro miembro de la familia.

Lista de Verificación Antes de Grabar (No utilizado)

- ☐ Pilas verificadas en el grabador principal (H1) y de respaldo (iPad)
- ☐ Tarjeta SD vacía; configuración WAV / 44.1 kHz confirmada
- ☐ Prueba de nivel de 10 segundos grabada y reproducida
- ☐ Dispositivo de respaldo en modo avión, grabación iniciada
- ☐ Protocolo impreso, lápiz, agua disponible
- ☐ Clave de participantes actualizada con esta sesión
- ☐ Espacio tranquilo y cómodo confirmado (preferencia del padre/madre: hogar, escuela, lugar comunitario)
- ☐ Formulario de consentimiento de grabación listo (copia impresa en el idioma preferido)

SD 112 Dual Language Program Evaluation

Stakeholder Interview Protocols

Paridad Education Consulting

Oak Terrace Elementary | Red Oak Elementary
Spring 2026

These protocols are designed to support the SD 112 Dual Language Evaluation by gathering rich qualitative data from administrators, teachers, and parents. Questions are aligned to the eight core evaluation questions established in the evaluation framework and draw on the six qualitative domains identified in the SD 112 Qualitative Data Tracker (v1). Interviewers should read probes aloud only

SD 112 Dual Language Program Evaluation

Student Focus Group — Field Protocol

Paridad Education Consulting • Spring 2026 • Protocol 4 (Grades 3–5, No Recording)
• 30 min

NO RECORDING. Student responses are captured by handwritten notes only. No audio or video recording of students is permitted. The note-taker is your evidence base — plan accordingly.

Intended for: Focus groups of 3–4 students in grades 3–5, drawn from the same grade or adjacent grades (e.g., 3rd–4th together, or 4th–5th together). All four DLP sites. Conducted in English (kids will code-switch naturally; that’s welcome and worth noting). Two adults required: one facilitator, one note-taker. The facilitator focuses on rapport, eye contact, and managing the group. The note-taker sits slightly to the side, writes verbatim where possible, and does not speak unless invited by the facilitator. This division of labor is what compensates for the absence of recording.

Roles — Facilitator and Note-Taker

Facilitator:

Asks questions, maintains eye contact, builds rapport, paces the conversation, redirects when one student dominates, draws out quiet voices. Does not write during the session except to glance at the protocol.

Note-taker:

SD 112 Dual Language Program Evaluation

Teacher Focus Group — Field Protocol

Paridad Education Consulting • Spring 2026 • Protocol 2 (30-min Group Format) • 6 questions

Intended for: 30-minute focus groups of three classroom teachers drawn from different grade levels at the same school. Groups should mix English-medium and Spanish-medium teachers where possible. Conducted in the group's preferred common language — typically English, with code-switching welcome.

Focus-group purpose: This is designed to surface what 1-on-1 interviews and classroom observations cannot — teacher-to-teacher dialogue about how the dual language program is implemented and experienced. Grade-level differences will emerge naturally as teachers describe their work. Classroom technique, assessment practice, cross-cultural pedagogy, and individual administrator-support experiences are captured through other instruments. Six questions, tight pacing.

Timing Discipline — Read Before Starting

At 30 minutes with three teachers, you have roughly 90 seconds per teacher per question. The time budgets on each question are strict — if one question runs long, the next one loses the equivalent minutes. Wear a watch. Prefer to cut a probe over cutting a teacher's response.

Time budget:

- Opening script + voice ID round: 5 min
- Q1 (program structure): 5 min

Appendix D Proposed Year-Long Biliteracy Assessment Schedule

A Proposed Year-Long Biliteracy Assessment Schools

FIGURE

D VI.3.8

By language and type of measure

Rows are paired by language: **English** **Español** A cell spanning both rows is shared by both languages.

LANGUAGE	ANNUAL / INTERIM MEASURES	COMMON MEASURES ACROSS CLASSROOMS	CLASSROOM-LEVEL DATA
English	FastBridge: fall screening WIDA Screener (new enrollees)	Home-language survey	Baseline routines established, such as journaling, student self-assessment and peer assessment
Español	FastBridge Español: fall screening	Home-language survey	Baseline assessment routines embedded in curriculum established as in English; setting a foundation for a student portrait
English	NWEA MAP: Reading (fall) NWEA MAP: Mathematics (fall)	Content-based oral language and written language samples and establishment of inter-rater reliability of interpretation with teachers using a common <u>rubric</u> or set of descriptors	Into Reading / CKLA Skills: weekly (as appropriate) and module embedded assessment
Español	NWEA MAP Lectura (fall) NWEA Mathematics in Spanish (fall)	Content-based oral language and written language samples and establishment of inter-rater reliability of interpretation with teachers using a common <u>rubric</u> or set of descriptors	iArriba la Lectura! / Caminos: weekly (as appropriate) and module embedded assessment
English	—	Writing sample 1: common prompt and <u>rubric</u>	Determination of how to capture <u>cross-linguistic transfer</u> (<u>translanguaging</u>); setting up the structure of an e-portfolio
Español	—	Writing sample 1: common prompt and <u>rubric</u>	
English	—	<u>Cross-linguistic transfer</u> inventory	Module assessments, including open-ended questions
Español	—	<u>Cross-linguistic transfer</u> inventory	Module assessments, including open-ended questions
English	NWEA MAP: Reading (winter) NWEA MAP: Mathematics (winter)	Writing sample 2: common prompt	Grade-level team interpretation of student samples, including <u>cross-linguistic transfer</u>
Español	NWEA MAP Lectura (winter) NWEA Mathematics in Spanish (winter)	Writing sample 2: common prompt	Grade-level team interpretation of student samples, including <u>cross-linguistic transfer</u>

English	WIDA ACCESS for English learners	—	Module assessments, including open-ended questions
Español	Spanish language proficiency measure for all students	—	Module assessments, including open-ended questions
English	WIDA ACCESS window	Content-based oral language sample 2	Student-led conferences; biliteracy portrait update
Español	Spanish language proficiency measure	Content-based oral language sample 2	
English	—	Writing sample 3: common prompt	Module assessments, including open-ended questions
Español	—	Writing sample 3: common prompt	Grade-level team interpretation of samples, including translanguaging
English	IAR : ELA and Mathematics (Gr 3–5)	—	Curricular module assessments, including open-ended questions
Español	Spanish language arts test, as appropriate	—	
English	FastBridge : (spring) NWEA MAP: Reading (spring) NWEA MAP: Mathematics (spring)	Writing sample 4: common prompt	Biliteracy portrait review with students
Español	FastBridge Español : (spring) NWEA MAP Lectura (spring) NWEA Mathematics in Spanish (spring)	Writing sample 4: common prompt	Biliteracy portrait review with students
English	Data archived to central database	Annual biliteracy assessment framework review	Student e-portfolio analyzed and archived
Español	Data archived to central database		

READING THE TABLE

Read across a language row to see how Spanish and English are treated across the types of assessment.

A common writing prompt with a shared interpretation session produces four samples a year that, once [inter-rater agreement](#) is established, can be aggregated and compared across classrooms as these measures offer a unique opportunity to examine [cross-linguistic transfer](#), a marker of [pillar 3, sociocultural competence](#). At that time, the value of [NWEA MAP](#) interim measures versus curriculum inspired oral language and written language samples should be determined.

ABOUT THE AUTHORS

PARIDAD ASSOCIATES

John F. Hilliard, MS Ed – Project Director

John F. Hilliard is Co-President and founder of Paridad Education Consulting. With more than 30 years of experience in bilingual and dual language education, John brings a comprehensive perspective that spans classroom teaching, university instruction, and national-level consultation. John was named Award-Winning Bilingual Teacher of the Year for New York State (1990-91). John is Co-Developer of Common Core and More! Summer Institute (CoCoMo) and the Dual Language Summer Institute (DUSI, 2007-2011). He is co-author and project director of DualU, a comprehensive professional learning curriculum for dual language program stakeholders. John is a National Consultant to WIDA Consortium's SALSA/PODER advisory panel for Spanish Language Development Standards, and co-author and developer, with Dr Margo Gottlieb, of EL ESPEJO®, a comprehensive curriculum framework for integrating language and content standards. John was a bilingual elementary teacher in Buffalo, NY for eight years and currently works in supporting dual language programs from PreK through middle school. John is co-author on various publications in the area of dual language and a frequent presenter and panelist in dual language conferences.

Cristina Sánchez-López, MS Ed – Co-Principal Investigator and On-Site Evaluation

Cristina Sánchez-López is Co-President at Paridad Education Consulting with 30 years of experience as a bilingual educator. Her expertise is in supporting the education of multilingual learners, particularly learners who experience challenges and students with disabilities. Cristina has extensive experience developing culturally and linguistically sustaining multi-tiered systems of support (MTSS) for multilingual learners and multilingual learners with disabilities. She supports educators in the areas of curriculum development, mathematics instruction, teaching literacy across content areas, and in implementing instructional approaches for multilingual learners with disabilities. Cristina facilitates professional learning communities and technical assistance nationally and internationally. She has served as an advisor to state departments of education in the US and ministries of education in Canada on issues related to the education of multilingual learners. Cristina has authored and co-authored various books, articles and book chapters on the education of multilingual learners.

COLLABORATING EXPERTS

Margo Gottlieb is Co-founder and Lead Developer of the WIDA Consortium and one of the nation's foremost experts on language development standards and assessment for multilingual learners. Over more than three decades, Margo led the creation of WIDA's language development frameworks for standards-referenced assessment systems (2004–2023), now used by 42 states and territories. She served on the U.S. Department of Education's inaugural National Technical Advisory Council (2008) and is a Fulbright Senior Specialist in Assessment and Evaluation. She also works extensively with states and districts nationally on assessment systems for multilingual learners in English and Spanish. She is the author of *Assessing Multilingual Learners: Bridges to Empowerment* (3rd ed., 2024), *Assessment in Multiple Languages: A Handbook for School and District Leaders* (2022), and *Classroom Assessment in Multiple Languages: A Handbook for Teachers* (2021), and is co-author, with Andrea Honigsfeld, of *Collaborative Assessment for Multilingual Learners and Their Teachers* (2025). Margo holds a PhD in Public Policy Analysis, Evaluation Research and Program Design from the University of Illinois-Chicago. In 2025, she was inducted into the inaugural class of the Multilingual Education Hall of Fame, and in 2016 she was recognized in TESOL International Association's 50@50.

Fred Genesee is Professor Emeritus at McGill University and one of the world's most respected scholars in bilingualism, dual language education, and second language acquisition. Over more than four decades, Fred's research has shaped the theoretical and empirical foundations of immersion and dual language programs internationally, including seminal studies of long-term outcomes in academic achievement, bilingual proficiency, and biliteracy development. His work spans bilingual first and second language acquisition and the relationship between language proficiency and academic achievement, with particular attention to outcomes for diverse learners, including English learners, students with learning difficulties, and students from varied linguistic and socioeconomic backgrounds. Fred brings deep expertise in interpreting

quantitative and longitudinal program data and has advised ministries of education, school districts, and educational organizations worldwide on bilingual education effectiveness and evidence-based program models. He holds a PhD in Psychology (Language and Cognitive Development) from McGill University. Fred is co-author of *Dual Language Instruction from A to Z* and *Educating English Learners*, has authored more than 200 peer-reviewed articles, book chapters, and research reports, and is a frequent keynote speaker at international conferences on dual language education and bilingualism.

Evaluator stance

Where evidence and professional judgment are used is marked where a recommendation relies on professional judgment in addition to collected data, which is legitimate and sometimes unavoidable, it is identified as such at the point of use.

Recommendations appear once, in Section VI. Section V reports findings only. This separation lets a reader accept a finding while disputing a recommendation, or the reverse.

Evaluator Independence Disclosure

Paridad Education Consulting holds no current contracts with District 112 and, to its knowledge, none of its proprietary frameworks are implemented in the current program. The Project Director advised the planning and professional development of the district's dual language program at its launch and District 112 appears among the Illinois districts the team has supported. This prior work is not a current contract. It is disclosed here for transparency.