



ARIZONA DEPARTMENT OF EDUCATION

Each reading intervention program will be scored using a rubric. A response must meet **all** the criteria in each scored area to be included on the K-3 MOWR vetted reading intervention program list. Peer Review will be noted but does not disqualify overall approval.

Study: [EVALUATING THE EFFECTIVENESS OF THE SIPPS PROGRAM](#)

Date of Study: January 2022 -21/22 SY

Grade Levels of Study: 1st and 2nd graders

Reading Intervention Program's Marketed Name: SIPPS Systematic Instruction in Phonological Awareness, Phonics, and Sight Words

	Rating or Measure Select: Y or N	Rationale
ESSA Evidence Tier	Y	Tier 2
Independent Researchers (Y/N)	Y	SEG Measurement
Established Measure (Y/N)	Y	Reading skills were measured using the DIBELS® Assessment. The DIBELS assessment was administered in September 2021 at the beginning of the school year (pretest) and again in the middle of the school year in January 2022 (posttest). DIBELS is a widely used measure of reading skills. The test publisher reports test reliability above .90 and provides validity support drawn from several studies. The DIBELS assessment includes several subtests, each reported on a subtest-specific vertical developmental scale bridging grades 1–12. This common scaling across grade levels supports the measurement of growth over time (regardless of grade). DIBELS also reports an aggregated composite score integrating the contributions of all subtests in a single score, on a scale ranging from 200 to 500.

		The DIBELS assessment includes the following subtests, each measuring a specific reading component skill. • LNF: Letter Naming Fluency • PSF: Phonemic Segmentation Fluency • NWF CLS: Nonsense Word Fluency–Correct Letter Sounds • NWF WRC: Nonsense Word Fluency–Words Recoded Correctly • CLS: Correct Letter Sounds • WRF: Word Reading Fluency ORF: Oral Reading Fluency • Maze: Identifying missing words in a passage • Composite Score: Aggregate score integrating the results for the specific tests above as applicable to the grade level
Sample Size	Y	626 students (313 treatment and 313 control) Pasco School District Pasco is a diverse district with approximately 48,000 students classified as White, 6,000 classified as Black, 20,000 classified as Hispanic, and 7,000 classified as “other.” About half (52%) of the students are eligible for free or reduced-price lunch The design of the study did not permit random sampling of participants, so matched groups of treatment and control groups were selected to ensure comparability of the groups at the outset of the study. Treatment and control group students were matched using propensity score matching; a logistic regression-based propensity score model was used. Students were matched on a composite propensity score based on initial reading ability (DIBELS pretest scores), gender, race, ethnicity, and economic status (free or reduced-price lunch eligibility). Sampling was done without replacement and matches were required to be within 1% of each other for the composite propensity score
Research Design	Y	Quasi-experimental- pre-post, treatment-control group design
Outcome	Y	Five hundred seventy students remained in the final analytic sample, after accounting for the attrition of 56 students (9%) who moved out of the district between October 2021 and January 2022 or otherwise were unavailable to continue the study. Twenty-eight students from the treatment group and 28 students from the control group were lost

		through attrition.
Effect Size	Y	The difference in reading skills growth (dependent variable) for the treatment and control groups (independent variable) was examined, adjusting for any differences in the initial reading ability of the groups at the outset of the study (covariates). Students receiving reading intervention that included instruction using SIPPS showed significantly greater growth in reading skills than did students receiving reading intervention without SIPPS instruction ($F = 10.72$, $df=1/570$; $p<.001$; Effect Size=.25). 16 Evaluating the Effectiveness of the SIPPS Program The treatment group achieved a mean reading skills posttest score of 364.30, while the control group achieved a mean reading skills posttest score of 359.46. This 4.84-point difference in mean scores reflects an effect size of .25. Effect size was computed as the difference in adjusted means divided by the total group unadjusted standard deviation.
*MTSS Level of Support	Y	Tier 2 and Tier 3
*Peer Review (Y/N)	N	

MOWR Committee findings:

-Will continue to be on the MOWR Intervention Program Vetted List (4th & 5th Edition).