Advancing Rural Computer Science

Final Technical Report:

Impact Study and Fidelity Study

August 28, 2025

Jennifer L. Maeng
Timothy Konold
Khushbu Singh

Table of Contents

Title and Abstract

Advancing Rural Computer Science

The study examined the Advancing Rural Computer Science (ARCS) program, a professional development initiative designed to improve elementary teachers' computer science (CS) content knowledge, pedagogical practices, and self-efficacy, with the goal of enhancing K–5 students' interest in and knowledge of CS.

The external evaluation employed a randomized controlled trial (RCT) design, with schools as the unit of random assignment. For two cohorts of schools, school teams were assigned to either the immediate treatment or delayed treatment (control) condition. The comparison condition (delayed treatment/control) was business-as-usual instruction, where teachers received no additional CS-focused professional development (PD) beyond standard district offerings. The program was implemented without substantive adaptations during the study.

The study was conducted in elementary schools located in across Virginia, with a focus on serving students in rural divisions and who are traditionally underrepresented in STEM and CS education, including Black, Hispanic, and mixed-race students, students from economically disadvantaged backgrounds, and those in geographically isolated communities. The sample included K–5 teachers and their students.

Teachers in schools randomized in the treatment condition completed a 5-day summer academy and academic follow up in year 1 and microcredentialing and academic year follow up in year 2 of the intervention and engaged in professional learning communities across the two years of the intervention. Teachers in schools randomized into the control condition were eligible to participate in the two-year intervention the year following data collection.

For both cohorts of schools, key Year 1 outcomes measured included teachers' CS content and pedagogical knowledge, CS self-efficacy, and instructional implementation frequency, and students' CS content knowledge and interest in CS.

Conclusions: Results indicate that for two cohorts of schools, one year of ARCS participation improved student CS content knowledge; controlling for all model covariates. The treatment effect indicated that treatment school student content knowledge means were .872 points greater than control group schools (γ_{oo} = 10.10); γ_{01} = .872, p = .0005, g = .186. However, there was no statistically significant improvement in student CS interest; controlling for all model covariates, the treatment effect indicated that treatment school means were .508 points greater than control group schools (γ_{oo} = 8.24); γ_{01} = .508, p = .057, g = .036.

Results indicated that for two cohorts of schools, one year of ARCS participation teachers' CS content knowledge and self-efficacy for teaching CS. Controlling for all model covariates, the treatment effect indicated that treatment school teacher content knowledge means were 1.50 points greater than control group schools (γ_{oo} = 9.04); γ_{01} = 1.50, p = .001, g = .513. Controlling for all model covariates, the treatment effect indicated that treatment school

teacher self-efficacy means were 8.11 points greater than control group schools (γ_{oo} = 9.25); γ_{01} = 8.11, p < .001, g = .221.

Background

A growing national demand for STEM and Computer Science (CS) proficiency highlights the urgent need to expand access to high-quality CS education, particularly in underserved and rural elementary schools. An estimated 2.4 million STEM and CS positions were unfilled in the U.S. workforce as of 2018, with projections indicating that over 50% of all jobs will soon require some level of STEM literacy (Carneval et al., 2013; Smithsonian, n.d.).

Despite this demand, access to CS education in K-12 schools remains uneven. Only 29% of rural schools offer CS education (Code.org, 2018), and students from underrepresented backgrounds such as those who are economically disadvantaged, from race and ethnic groups traditionally underrepresented in STEM, or located in rural areas face systemic barriers to early exposure and sustained engagement in CS learning (National Science and Technology Council, 2018).

The root causes of this inequity are multifaceted. Teachers in rural and high-need districts often lack access to content-rich, sustained professional development (PD) opportunities that build CS content knowledge and pedagogical skill (Abell, 2007; Supovitz & Turner, 2000). In addition, early indicators of STEM persistence such as interest, confidence, and self-efficacy begin to decline in elementary school for underrepresented students, particularly Black, Hispanic, female, and rural learners (Dickerson et al., 2014; Ireland et al., 2018; Connors-Kellgren et al., 2016). Addressing these root causes requires interventions that begin at the elementary level, integrate CS into core content areas, and use inclusive, culturally responsive teaching practices that reflect students' experiences and communities (Brown-Jeffy & Cooper, 2011; Margolis et al., 2008).

The ARCS intervention builds on a foundation of evidence-based practices. Prior studies have demonstrated that PD programs combining content knowledge development with sustained collaborative learning such as summer institutes followed by school-year support led to improved teacher confidence, knowledge, and instructional practice (Meyers et al., 2015; Maeng et al., 2015; Taylor et al., 2017). Further, integrating CS with core content areas such as math, reading, and science increases both the feasibility and effectiveness of implementation in elementary classrooms (Sáez-López et al., 2016). Microcredentialing has also emerged as a promising strategy for recognizing and motivating teacher growth in specialized content areas (Crow & Pipkin, 2017; DeMonte, 2017).

This body of research supports the design of a professional development model that is blended, sustained, and personalized. The ARCS intervention aims to address both teacher and student needs by providing teachers with professional learning experiences such as a summer academy taught by CodeVA and a microcredential pathways aligned to state CS standards, while supporting them in developing and implementing integrated, culturally responsive CS lessons.

This approach is particularly timely in Virginia, where the adoption of K–12 CS Standards of Learning and enabling legislation for microcredentials has created a policy environment receptive to innovation and scale. The proposed work contributes to the field by advancing a scalable, evidence-based model that addresses equity in early CS education and supports sustainable improvements in teaching and learning in underserved rural contexts.

Impact Study

Study Description

Research Questions for the Study

The impact evaluation of the ARCS program addressed the following research questions through a randomized controlled trial design, in which the intervention group was compared to a business-as-usual control group after Year 1 of the ARCS intervention in elementary schools in Virginia.

Confirmatory

- What is the effect of ARCS participation on school-level CS content knowledge among students in grades 3–5, compared to students in schools assigned to the business-as-usual condition, as measured by the CKACS (Computer Knowledge and Attitudes toward Computer Science) assessment administered in Fall and Spring?
- What is the effect of participation in the ARCS program, compared to business-as-usual instruction, on the school-level average of 3–5 students' interest in computer science (CS) as measured by pre- and post-intervention student surveys?

Exploratory

- What is the effect of ARCS professional development on K-5 teachers' CS content knowledge and self-efficacy, compared to teachers in business-as-usual condition, as measured by pre- and post-intervention teacher surveys and assessments?

These research questions were designed to assess both student- and teacher-level outcomes, with attention to outcome domains including CS interest, content knowledge, and teacher self-efficacy. Measures were collected at multiple time points including pre- and post-intervention phases and end-of-year follow-ups, depending on cohort assignment and exposure to professional development activities.

Intervention Condition

The program under study is the ARCS initiative, a PD program developed by Old Dominion University in partnership with Code VA and the Virginia Department of Education. The primary aim of ARCS is to improve elementary students' CS content knowledge and interest by strengthening K–5 teachers' CS content knowledge, pedagogical skills, and instructional confidence. The program specifically targeted teachers and students in rural Virginia, with an emphasis on serving populations historically underrepresented in STEM fields.

The ARCS program consists of the following core components:

Summer Professional Development Academy (Year 1)

- Services: Virtual PD workshops (synchronous and asynchronous).
- Instructional Approaches: Direct instruction, modeling of CS-integrated lessons, hands-on coding and unplugged activities, discussion, reflection, and collaborative planning.
- Session Topics: Core CS concepts (e.g., algorithms, programming, data, and networks), strategies for interdisciplinary integration (e.g., CS in literacy and math), culturally responsive instruction, and equity in CS education.
- Target: Build foundational teacher capacity in CS content and instructional integration.

Academic Year Follow Up

- Services: Networked Improvement Community (NIC) / Learning Bytes Webinars (Year 1 & 2), Classroom Resources, Coaching
- Instructional Approaches: Monthly virtual meetings, peer collaboration, reflective practice, sharing implementation strategies, individual and small group coaching.
- Topics: Practical challenges in implementation, adapting CS content to classroom needs, sharing successes and failures, instruction around how to use resources (e.g, Spheros).
- Target: Support classroom implementation of CS-integrated lessons and foster a professional learning community.

Microcredentialing (Year 2)

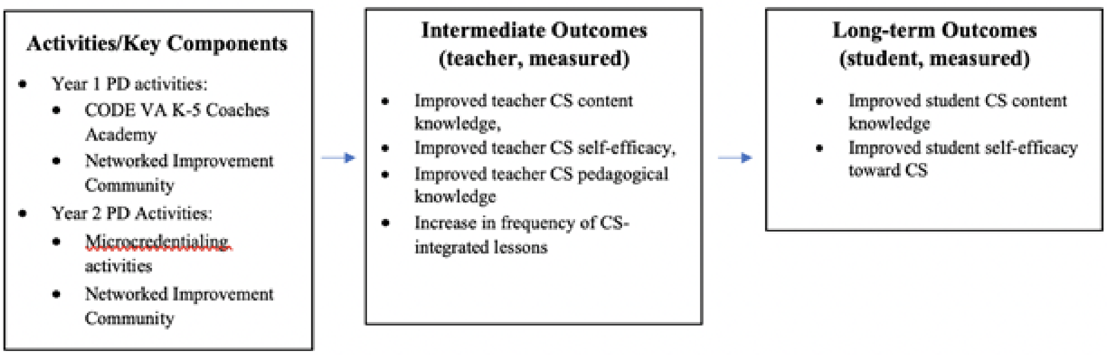
- Services: Canvas-based asynchronous instructional modules.
- Instructional Approaches: Asynchronous instructional models, Structured portfolio development for teachers to demonstrate CS instruction, feedback from coaches, self-assessment.
- Topics: (1) Introduction to Computer Science Principles, Digital Impact, and Digital Citizenship; (2) Computing Systems, Networks and the Internet, and Cybersecurity; (3) Algorithms and Programming, (4) Data and Analysis; and (5) Lesson Integration.
- Target: Validate teacher growth and sustained integration of CS into teaching practice.

Students in participating teachers' classrooms engaged in CS content and activities, supporting the program's long-term goals of improving student interest and knowledge in computer science.

The intervention targeted both teacher-level outcomes (knowledge, pedagogy, self-efficacy, and instructional frequency) and student-level outcomes (CS content knowledge and interest), with a strong focus on equity and rural education.

Adaptations to the Program Model. No substantive adaptations were made to the core content or structure of the ARCS program for the study. However, minor delivery adjustments were implemented in response to logistical and contextual needs, including virtual delivery of Code VA Coaches Academy (proposed in the original grant proposal as in-person workshops) due to COVID-19 restrictions. Asynchronous materials and recorded sessions were used to accommodate teacher schedules. These minor adaptations-maintained fidelity to the program's core goals and learning objectives.

ARCS Logic Model for EIR Evaluation



Intensity/Duration of the Key Components

- Summer PD Academy: Approximately 30 hours over one week in Year 1 of the intervention.
- Networked Improvement Community (NIC)/Learning Bytes Webinar: Monthly sessions throughout Years 1 and 2 (6–10 sessions per year).
- Microcredentialing: Year 2; flexible self-pacing over several months depending on teacher schedule.
- Instructional Implementation: Ongoing throughout the academic year as teachers embed CS into their curriculum.

Methods of Delivery of Key Components

- Summer Code VA Coaches Academy: Delivered virtually, led by facilitators from Code VA and ODU.
- NIC/Learning Bytes: Delivered virtually via online platforms (e.g., Zoom, discussion forums) to enable statewide collaboration led by facilitators at Code VA and ODU.
- Microcredentialing: Delivered through an online portfolio platform (Canvas) with support from coaches and peer reviewers from ODU.
- Instructional Implementation: Delivered by teachers in their own classrooms, using lesson plans developed during PD and supported by the ARCS framework.

Program or Service Implementation

The ARCS intervention spanned two years for each of two cohorts of teachers, including a 5-day Coaches Academy and ongoing school-year professional development delivered through Networked Improvement Communities (NICs) and Learning Bytes webinars. The delivery format was virtual (synchronous and asynchronous) group-based learning, with additional coaching and peer collaboration throughout the year.
The program was delivered by experienced educators and facilitators from CodeVA, the Virginia Department of Education, and faculty from Old Dominion University. The intervention took place outside of regular school hours and was designed to build the teacher capacity to integrate computer science into the elementary curriculum.

Fidelity to the program was measured through multiple methods, including attendance and participation records and teacher surveys.

Setting

Location: Virginia, United States- South

Setting Type: Public Elementary Schools

School Environment: In school, regular classrooms

Geographic Focus: Primarily rural and semi-rural school districts

Student Population: grades 3-5 students in Virginia public elementary schools, including Title 1, economically disadvantaged, and students from backgrounds traditionally underrepresented in STEM (e.g., female, Black, Hispanic)

Comparison Condition

The comparison condition (control/delayed treatment) in the ARCS study was defined as a business-as-usual model, which qualifies as a treatment-as-usual control group. Teachers assigned to this condition did not receive the ARCS PD during the first year after randomization (they engaged in data collection only as a control group) but were offered the ARCS PD in the year following their control year, thus functioning as a delayed intervention or waitlist control group. This ensured that all participants eventually received the intervention, though the control group served as a year 1 comparison.

The same outcome measures (i.e., CS content knowledge assessment, interest survey) were administered to both treatment and comparison groups of students at two timepoints (pre/year 1 end) that were consistent across groups. For teachers, the same outcome measures (i.e., CS content knowledge assessment, self-efficacy survey, implementation survey) were administered to both treatment and comparison groups of students at multiple timepoints (pre/mid-year, year 1 end) that were consistent across groups.

Comparison group teachers did not participate in the CodeVA Coaches Academy during or prior to their control year.

Study Participants

Identification/selection of study districts
The sample was drawn from the district partners who signed on as part of the application and other districts that were identified by Virginia Department of Education as rural (59% or 78 out of 132 school divisions). K-5 teachers from identified rural schools (defined based on the VDOE list of rural divisions) were included in one of two cohorts in the Year 1 impact study. Combined, these divisions had 95 elementary schools with 3,495 K-5 teachers and 43,292 students. Applications to participate in the ARCS PD were submitted at the school level.

Identification/selection of study schools
Rural divisions in Virginia included 95 elementary schools with 3,495 K-5 teachers and 43,292 students. Schools in these districts were recruited to apply to participate in the ARCS PD. Identification and selection of study schools were based on those that applied to participate in the ARCS PD. Schools selected a grade level in which all students would participate. The evaluation team randomized schools at a ratio of 50% to treatment and control conditions for each of the two cohorts. Each cohort was randomized at the school level. Approximately 100 schools were included in the analytic sample.

Identification/selection of study teachers
All teachers in school districts meeting the criteria identified above were eligible to apply and participate in the study. The number of teachers expected to participate in the 100 schools in the analytic sample across two cohorts was 440. Teachers in the analytic sample were members of the school-level cluster prior to randomization.

Identification/selection of students
All students within a school’s focal grade (in that school’s targeted grade band in which all teachers participated) in randomized schools were included in the impact analysis. Students were members of the schools prior to randomization. Students who joined the schools after randomization were not included.

The analysis treated students as nested within schools rather than within teachers (a two-level model in which variation at the teacher level was aggregated to the school level). The target number of students across the two cohorts of schools was 2,500. As described in the section above, random assignment occurred at the school level with no stratification. The evaluation team conducted random assignment of schools to treatment and control groups. For each teacher, a roster of students assigned to the class was collected in the fall of the school year so student-level attrition could be documented. Parent notification (passive consent) was required.

Sample Alignment with Those Served by the Program

The evaluation sample was based on a RCT in which schools were the unit of randomization, and all teachers within selected schools were assigned to either the treatment or comparison condition. Students in the classrooms of participating teachers are included in the evaluation.

Design and Measures

Independence of the Impact Evaluation

The impact evaluation of the ARCS program was conducted independently by researchers at the University of Virginia, who were not involved in the development or implementation of the intervention. The intervention itself was developed and implemented by Old Dominion University, in collaboration with CODE VA and the Virginia Department of Education (VDOE).

Key evaluation activities conducted by UVa independently of the implementation team included the random assignment of schools, collection of outcome data, and all impact analyses. These were conducted solely by UVa researchers. The evaluators' institutional affiliation, distinct from that of the intervention developers, affirms the independence of the evaluation. UVa was engaged to provide an objective and unbiased assessment of the ARCS program's impact on teacher and student outcomes.

Pre-registration of the Study Design

This study was pre-registered in the Registry of Efficacy and Effectiveness Studies (REES), Registry ID: 7200.1v1. The registered protocol included two confirmatory research questions related to the effects of the ARCS professional development program on (1) grade 3–5 student computer science content knowledge, and (2) student interest in computer science after one year, compared to a business-as-usual condition.

Design

ARCS used a RCT design. Randomization occurred at the school level, schools that had teachers apply to ARCS were randomized into either the treatment (early start) or control (delayed start) condition at the school level; all teachers in a school were assigned the same condition. Teachers in schools randomized into the treatment immediately received a 2-year professional development (PD), while teachers in schools randomized into the control condition began the PD a year later. Ultimately, all teachers received the PD. This process was repeated for two cohorts of schools, one was randomized in Spring 2021 and one cohort was randomized in Spring 2023.

Teachers were recruited through the end of April. Randomization occurred at the school level twice, in early March and in early May, and schools were notified of their assignment in mid-May.

ODU and UVA both had responsibility for recruitment and obtaining consent; however, UVA was responsible for random assignments and impact study data collection and analysis.

Regardless of condition, teachers completed CS-content knowledge, CS-pedagogical knowledge, CS-self-efficacy, and CS-frequency of implementation assessments at multiple timepoints throughout the two years of the study (three for control teachers). Teachers completed a pre and post-PD Perceptions survey and a year-end survey each year of participation and they completed a mid-year implementation survey.

Each year of teacher participation, students in teachers' classes completed CS-content knowledge and interest assessments during the first four weeks of school and during April (end of the academic year).

Measures

Student Baseline and Outcome Measures

Student CS Content Knowledge Assessment: Grades 3-5 students in participating schools completed a computer science performance assessment twice each year. All students in a grade were assessed and the grade of students assessed was based on the teachers that applied from that school (i.e., if 3rd grade teachers in the school applied, then 3rd grade students were assessed. In instances where teachers of multiple grades applied from a school, the evaluation team assigned the grade of students to be assessed). This assessment measured students' knowledge and understanding of integrative computer science as they analyze and solve complex problems. This measure meets the WWC Technology and Engineering Literacy outcome domain. Assessments were completed online. Assessments were scored by the external evaluator using a detailed rubric designed by the external evaluator after establishing scoring reliability. An overall score for each student was calculated. The student performance assessment was administered at the beginning and end of each school year for students in treatment and control teachers' classrooms with the end-of-year score used as the outcome variable and beginning of the year score used as baseline. Baseline and outcome scores will be averaged across participating students to obtain a school-level mean score content knowledge.

Student CS Interest Assessment: Grade 3-5 students completed a measure of attitudes toward computer science. All students in a grade were assessed and the grade of students assessed was based on the teachers that applied from that school (i.e., if 3rd grade teachers in the school applied, then 3rd grade students were assessed. In instances where teachers of multiple grades applied from a school, the evaluation team assigned the grade of students to be assessed). The instrument was developed from existing validated instruments. Items were be adapted to language appropriate for elementary students, and the evaluators established, internal consistency, and support for face and content validity. The interest assessment was administered at the beginning and end of each school year for students in treatment and control teachers' classrooms with the end-of-year score used as the outcome variable and beginning of the year score used as baseline. Baseline and outcome scores will be averaged across participating students to obtain a school-level mean score CS interest.

These measures were implemented with students in both cohorts of schools.

Teacher Baseline and Outcome Measures

Teacher CS Content Knowledge: This measure consists of 5 open-ended response items developed by the external evaluator with support for face validity established through expert review. Teacher responses will be coded as not aligned (0 point), partially aligned (1 point), and fully aligned (2 points) using a rubric developed by the external evaluator and an overall score (ranging from 0 to 10) calculated for content knowledge.

Teacher CS Self-efficacy: This measure consists of 9 Likert scale items adapted from the Teachers' Self-efficacy in Computational Thinking (Bean, Weese, Feldhausen, & Bell, 2015); a = .935) instrument. Modifications include using a 6-point scale instead of a 5-point scale, and replacing items 9 and 10, which relate to the Common Core and NGSS with a single item about the Virginia Standards of Learning. Cronbach's for the revised instrument was calculated using pilot data and determined to be .92 at pre-test and .92 at post-test, indicating good reliability. An overall score will be calculated for CS self-efficacy.

These measures were administered as a survey to treatment and control teachers before the first summer PD to establish baseline measures of each construct, at the end of the school year in year 1, and at the end of year 2. Year 1 end scores were used as outcome scores for exploratory analysis. Baseline and outcome scores were averaged across participating teachers to obtain a school-level mean score for content knowledge and self-efficacy.

For all analyses, potential school-level covariates included: school size, percent students receiving free and reduced priced meals, percent non-White students. Potential teacher-level covariates (aggregated to the school level) included: years of teaching experience, baseline score for content knowledge, pedagogical knowledge, and self-efficacy. Potential student-level covariates include baseline score for content knowledge and interest.

Some of these outcomes do not fall within WWC reviewable outcome domains; however, they can be used to answer critical questions in the field and for the grantee.

Sample Sizes and Attrition

This study used a cluster randomized controlled trial (RCT) design, with schools as the unit of randomization. The two cohorts of schools were combined for the year 1 impact analysis. In total, 68 schools were randomized to either the treatment condition (n = 35) or the comparison condition (control/delayed treatment) (n = 33) across two cohorts. At the time of randomization, the total number of participating students was 4,347, with 1,972 students in the treatment group and 2,375 in the comparison group (control/delayed treatment).

School- Level Attrition

Across both cohorts, 14 schools (9 treatment, 7 comparison) did not submit student post-test data, resulting in an overall cluster-level attrition rate of approximately 20.6%. Attrition was higher in the treatment group (22.9%) than in the comparison group (18.2%), yielding a differential attrition rate of 4.7 percentage points. According to What Works Clearinghouse Standards, with overall attrition of 20.6%, differential attrition is acceptable if it is below 5.3 percentage points under cautious standards and 9.9 under optimistic ones. The observed differential attrition (4.7 percentage points) falls below both thresholds, indicating low risk of bias due to attrition.

Student-Level Attrition

Overall, student attrition across both cohorts was 32.5%. Attrition was similar between the treatment (31.8%) and control (33.0%) conditions, resulting in a differential attrition rate of -1.2 percentage points. According to What Works Clearinghouse (WWC), for a 32% overall attrition rate, the cutoff for acceptable differential attrition is 3.8 percentage points (cautious) and 7.8 percentage points (optimistic). The observed differential of 1.2 percentage points is below both thresholds, indicating low risk of bias.

Analytic Sample

The final analytic sample for student outcomes included 1,591 students from the comparison group and 1,344 students from the treatment group who submitted complete post-test (year-end) data.

Table 2. Sample Sizes at Randomization and in Analytic Sample Needed to Assess Attrition for an RCT with Cluster-Level Assignment

Outcome Measure	Comparison Group				Treatment Group			
	Clusters[a]		Students[b]		Clusters[a]		Students[b]	
	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample
Cohort 1	11	9	1091	736	12	8	786	407
Cohort 2	22	18	1284	855	23	19	1186	937
Both Cohorts	33	27	2375	1591	35	27	1972	1344

[a] Clusters refer to schools that are randomized and retained in the analytical sample

[b] Students in the analytical sample are those from non-attritted clusters with post-test data

Teacher-Level Attrition

Overall, teacher attrition across both cohorts was 37.2%. Attrition for treatment (39.8%) and control (35.0%) conditions, resulting in a differential attrition rate of 4.8 percentage points According to the What Works Clearinghouse (WWC), for a 37% overall attrition rate, the cutoff for acceptable differential attrition is 4.2 percentage points (cautious) and 8.2 percentage points (optimistic). The observed differential of 4.8 percentage points falls between these thresholds, indicating moderate risk of bias.

Analytic Sample

The final analytic sample for teacher outcomes included 100 teachers from the comparison group and 88 from the treatment group who were in schools with post-test (year-end) data available.

Table 2. Sample Sizes at Randomization and in Analytic Sample Needed to Assess Attrition for an RCT with Cluster-Level Assignment.

Outcome Measure	Comparison Group				Treatment Group			
	Clusters[a]		Teachers[b]		Clusters[a]		Teachers[b]	
	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample	# Random-ized	# Analytic Sample
Cohort 1	12	10	43	25	11	10	32	17
Cohort 2	22	18	57	40	23	20	56	36
Both Cohorts	34	28	100	65	34	30	88	53

[a] Clusters refer to schools that are randomized and retained in the analytical sample

[b] Teachers in the analytical sample are those from non-attritted clusters with post-test data

Data Analysis and Findings

Baseline Equivalence

Baseline equivalence testing was not conducted for either the student or teacher analytic samples because overall and differential attrition were within acceptable thresholds for impact outcomes.

Program Effects: Student Outcomes (Confirmatory)

Note. This is a two-level model in which students are nested within schools. Random assignment occurred at the level of school, and attrition analysis at the school and student level resulted in low attrition at both levels.

Model equations for the primary/confirmatory research questions:

What is the impact of ARCS professional development (PD) on the content knowledge (and interest) of their students? The same modeling strategy was adopted for outcomes considered to be exploratory.

Level 1 Model: Student Level

$$Y_{ij} = B_{oj} + B_{1j}(X1_{ij}) + B_{2j}(X2_{ij}) + r_{ij}$$

Where;

Y_{ij} = student posttest (e.g., content knowledge) ,

B_{oj} = conditional mean of outcome score controlling for pretest,

B_{1j} = average within school pretest-posttest relationship,

$X1_{ij}$ = pretest score (e.g., content knowledge) for student i in school j, missing values represented by -99,

$X2_{ij}$ = dummy coded variable to reflect presence (= 0) or missing (= 1) pretest score ,

r_{ij} = random effect representing the difference between student ij's score and the predicted mean score for school j, r_{ij}~ND(0, σ^2).

Level 2 Model: School Teams

$$\beta_{oj} = \gamma_{oo} + \gamma_{01}W_j + \sum_q^Q \gamma_{oq}W_{qj} + U_{oj}$$

Where;

γ_{oo} = conditional posttest score for control schools,

γ_{01} = treatment effect (i.e., the conditional mean difference between treatment and control schools),

W_j = 1 if school j is an intervention school, and 0 if control,

W_{qj} = Q additional school level covariates (e.g., school size, %disadvantaged, %White, randomization round (measured with 4 dummy coded variables to represent the 5 randomization occasions),

γ_{oq} = Q coefficients corresponding to additional school level covariates,

U_{oj} = deviation of school j's mean from the grand mean, conditional on covariates, U_{oj}~ND(0, τo^2)

All models included random intercepts and were estimated with REML to avoid underestimating variance components that can occur with ML in instances of low cluster numbers relative to the number of fixed effects being estimated (i.e., model complexity). Level 1 variables were not cluster mean centered in order to investigate between group post-test differences while controlling for student level pretest scores (i.e., ML adjusted means as outcome models; see for example, Enders & Tofighi, 2007). Missing pretest scores in our RCT with low attrition were handled through the dummy variable approach (Puma et al., 2009) as advocated by WWC. Here, missing pretest scores were represented by -99 values, and an addition dummy coded indicator was added to the model to represent missing (= 1) and non-missing (= 0) pretest values. Hedges g effect sizes were calculated as the ratio of the treatment effect to the total variance:

$$g = \frac{\gamma_{01}}{\sigma^2 + t^2}$$

Results

Content Knowledge

An unconditional model indicated that 27.3% of the total post-test score content knowledge was attributed to differences between schools (ICC = .273). Contrasts between treatment and control schools on their average pretest content knowledge scores were non-significant (γ_{01} = -.0099, p = .96), after controlling for missingness as described above. However, we retained this covariate in our impact analyses to improve precision in our estimate of the treatment effect. Controlling for all model covariates, the treatment effect indicated that treatment school means were .872 points greater than control group schools (γ_{oo} = 10.10); γ_{01} = .872, p = .0005, g = .186.

Random effects:

Groups	Name	Variance	Std. Dev.
SCHID	(Intercept)	0.6189	0.7867
	Residual	4.0693	2.0172

Number of observations: 2934

Groups (SCHID): 54

Fixed effects:

Predictor	Estimate	Std. Error	df	t-value	p-value
Intercept	10.1003896	1.1371	65.7008903	8.882	< .001 ***
Condition	0.8717716	0.2355	53.8917272	3.702	< .001 ***
PRECK	0.3704022	0.0241	2933.4474958	15.379	< .001 ***
MpreCK	41.1153807	2.7079125	2933.4873271	15.183	< .001 ***
ENROLL	0.0007656	0.0007574	50.8609723	1.011	0.317
PDisadv	-2.6252544	1.0424768	54.8175854	-2.518	0.042*
PWhite	0.2855877	0.5815706	58.5473281	0.491	0.625
RAND1	-1.1122893	0.4568264	55.2935571	-2.435	0.018*
RAND2	-1.2502061	0.5565386	52.8987063	-2.246	0.029*
RAND3	0.0927688	0.3643955	53.2640196	-0.255	0.800
RAND4	1.3660929	0.5328147	53.1849955	2.564	0.013*

Interest

An unconditional model indicated that 5.5% of the total post-test interest score was attributed to differences between schools (ICC = .055). Contrasts between treatment and control schools on their average pretest interest scores were non-significant (γ_{01} = -.10, p = .49), after controlling for missingness as described above. However, we retained this covariate in our impact analyses to improve precision in our estimate of the treatment effect. Controlling for all model covariates, the treatment effect indicated that treatment school means were .508 points greater than control group schools (γ_{oo} = 8.24); γ_{01} = .508, p = .057, g = .036.

Random effects:

Groups	Name	Variance	Std. Dev.
SCHID	(Intercept)	0.5577	0.7468
	Residual	13.6956	3.7008

Number of observations: 2919
Groups (SCHID): 54

Fixed effects:

Predictor	Estimate	Std. Error	df	t-value	p-value
Intercept	8.2360796	1.2336083	49.1784734	6.676	< .001 ***
Condition	0.5084085	0.2597931	44.9444511	1.957	0.057
PREINT	0.3969241	0.0255365	2902.6052290	15.543	< .001 ***
MpreINT	44.1324077	2.8645472	2901.7524393	15.406	< .001 ***
ENROLL	0.0008225	0.0008168	39.7882559	1.007	0.320
PDisadv	-1.1463738	1.1422632	41.8996035	-1.004	0.321
PWhite	-0.7067995	0.6568014	50.3819488	-1.076	0.286
RAND1	-0.2906409	0.5091247	48.0975723	-0.571	0.571
RAND2	-0.3043791	0.6058660	40.9420770	-0.502	0.618
RAND3	0.0959035	0.4022221	47.0288510	-0.238	0.813
RAND4	-0.2357935	0.5885296	47.1750750	-0.401	0.690

Confirmatory Impact Analysis Results (Cluster-Level Assignment Study)

Outcome Measure	Comparison Group				Treatment Group				Treatment – Control Difference	Standard Error	Standardized Difference	p-value
	Sample Size		Mean	Standard Deviation	Sample Size		Model-Adj. Mean	Standard Deviation				
	# Clusters	# Students			# Clusters	# Students						
Content knowledge	27	1590	10.10	2.30	27	1344	10.97	2.54	0.872	.24	0.19	< .001
Interest	27	1585	8.24	3.96	27	1334	8.24	3.92	0.508	.26	0.04	=0.057

Program Effects: Teacher Outcomes (Exploratory)

Note. This is a two-level model in which teachers are nested within schools. Random assignment occurred at the level of school, and attrition analysis at the school and teacher level resulted in low attrition at both levels.

Model equations for research questions:

What is the impact of ARCS professional development (PD) on the content knowledge (and self-efficacy) of teachers?

Level 1 Model: Teacher Level

$$Y_{ij} = B_{oj} + B_{1j}(X1_{ij}) + B_{2j}(X2_{ij}) + r_{ij}$$

Where;

Y_{ij} = teacher posttest (e.g., content knowledge) ,

B_{oj} = conditional mean of outcome score controlling for pretest,

B_{1j} = average within school pretest-posttest relationship,

$X1_{ij}$ = pretest score (e.g., content knowledge) for teacher i in school j, missing values represented by -99,

$X2_{ij}$ = dummy coded variable to reflect presence (= 0) or missing (= 1) pretest score ,

r_{ij} = random effect representing the difference between teacher ij's score and the predicted mean score for school j, r_{ij}~ND(0, σ^2).

Level 2 Model: School Teams

$$\beta_{oj} = \gamma_{oo} + \gamma_{01}W_j + \sum_q^Q \gamma_{oq}W_{qj} + U_{oj}$$

Where;

γ_{oo} = conditional posttest score for control schools,

γ_{01} = treatment effect (i.e., the conditional mean difference between treatment and control schools),

W_j = 1 if school j is an intervention school, and 0 if control,

W_{qj} = Q additional teacher level covariates (e.g., female (=1, other = 0), White (=1, other = 0), Hispanic (=1, other = 0), years experience, degree in education (= 1, other = 0), STEM degree (=1, other = 0), amount of other professional development in past 3 years, randomization round (measured with 4 dummy coded variables to represent the 5 randomization occasions),

γ_{oq} = Q coefficients corresponding to additional school level covariates,

U_{oj} = deviation of school j's mean from the grand mean, conditional on covariates, U_{oj}~ND(0, τo^2)

All models included random intercepts and were estimated with REML to avoid underestimating variance components that can occur with ML in instances of low cluster numbers relative to the number of fixed effects being estimated (i.e., model complexity). Level 1 variables were not cluster mean centered in order to investigate between group post-test differences while controlling for student level pretest scores (i.e., ML adjusted means as outcome models; see for example, Enders &

Tofighi, 2007). Missing pretest scores in our RCT with low attrition were handled through the dummy variable approach (Puma et al., 2009) as advocated by WWC. Here, missing pretest scores were represented by -99 values, and an addition dummy coded indicator was added to the model to represent missing (= 1) and non-missing (= 0) pretest values. Hedges g effect sizes were calculated as the ratio of the treatment effect to the total variance:

$$g = \frac{\gamma_{01}}{\sigma^2 + t^2}$$

Results

Content Knowledge

An unconditional model indicated that 30.9% of the total post-test score content knowledge was attributed to differences between schools (ICC = .309). Contrasts between treatment and control schools on their average pretest content knowledge scores were non-significant (γ_{01} = -.702, p = .71), after controlling for missingness as described above. However, we retained this covariate in our impact analyses to improve precision in our estimate of the treatment effect. Controlling for all model covariates, the treatment effect indicated that treatment school means were 1.50 points greater than control group schools (γ_{oo} = 9.04); γ_{01} = 1.50, p = .001, g = .513.

Self-Efficacy

An unconditional model indicated that 5.5% of the total post-test interest score was attributed to differences between schools (ICC = .055). Contrasts between treatment and control schools on their average pretest interest scores were non-significant (γ_{01} = -.10, p = .49), after controlling for missingness as described above. However, we retained this covariate in our impact analyses to improve precision in our estimate of the treatment effect. Controlling for all model covariates, the treatment effect indicated that treatment school means were 8.11 points greater than control group schools (γ_{oo} = 9.25); γ_{01} = 8.11, p < .001, g = .221.

Exploratory Impact Analysis Results (Cluster-Level Assignment)

Outcome Measure	Comparison Group: Sample Size: # Clusters	Comparison Group: Sample Size: # teachers	Comparison Group: Mean	Comparison Group: Standard Deviation	Treatment Group: Sample Size: # Clusters	Treatment Group: Sample Size: # teachers	Treatment Group: Model-Adj. Mean	Treatment Group: Standard Deviation	Treatment – Control Difference	Standard Error	Standardized Difference	p-value
Content knowledge	28	58	9.04	1.96	29	50	10.54	1.80	1.50	.407	.51	< .001
Self-efficacy	28	61	9.25	9.65	30	53	17.36	6.99	8.11	1.28	.22	< .000001

Discussion

The study examined the Advancing Rural Computer Science (ARCS) program, a professional development (PD) initiative designed to improve elementary teachers' computer science (CS) content knowledge, pedagogical practices, and self-efficacy, with the goal of enhancing K–5 students' interest in and knowledge of CS. The external evaluation employed a randomized controlled trial (RCT) design, with schools as the unit of random assignment. For two cohorts of schools, school teams were assigned to either the immediate treatment or delayed treatment (control) condition. The comparison condition (delayed treatment/control) was business-as-usual instruction, where teachers received no additional CS-focused professional development beyond standard district offerings. The program was implemented without substantive adaptations during the study.

Confirmatory research questions related to student CS knowledge and interest in CS. Results indicated on average students in schools randomized into the treatment condition had higher CS content knowledge compared to students in schools randomized into the control condition; the treatment effect indicated that treatment school student content knowledge means were .872 points greater than control group schools (γ_{oo} = 10.10); γ_{01} = .872, p = .0005, g = .186. However, there was no statistically significant improvement in student CS interest; controlling for all model covariates, the treatment effect indicated that treatment school means were .508 points greater than control group schools (γ_{oo} = 8.24); γ_{01} = .508, p = .057, g = .036.

Exploratory research questions related to teacher CS knowledge and self-efficacy. Results indicated on average, after one year of the ARCS intervention, teachers in schools randomized into the treatment condition had greater CS content knowledge and self-efficacy compared to teachers in schools randomized into the control condition. Controlling for all model covariates, the treatment effect indicated that treatment school teacher content knowledge means were 1.50 points greater than control group schools (γ_{oo} = 9.04); γ_{01} = 1.50, p = .001, g = .513. Controlling for all model covariates, the treatment effect indicated that treatment school teacher self-efficacy means were 8.11 points greater than control group schools (γ_{oo} = 9.25); γ_{01} = 8.11, p < .001, g = .221.

Teachers in schools randomized in the treatment condition completed a 5-day summer academy and academic follow up in year 1 and microcredentialing and academic year follow up in year 2 of the intervention and engaged in professional learning communities across the two years of the intervention. Teachers in schools randomized into the control condition were eligible to participate in the two-year intervention the year following data collection.

Overall, these results indicate that both students and teachers in schools receiving one year of the ARCS intervention improved their CS content knowledge compared to students and teachers in schools not receiving the ARCS intervention. In addition, teachers in schools receiving one year of the ARCS intervention improved their CS self-efficacy compared to teachers in schools not receiving the ARCS intervention. Importantly, the microcredentialling year of the ARCS intervention (year 2) was not assessed because of the lack of a comparison group.

The observed findings from the ARCS evaluation align closely with the root causes and mechanisms identified in the literature (e.g., Abell, 2007; Supovitz & Turner, 2000). The intervention used a blended, sustained PD model to address key barriers such as limited access to sustained, content-rich professional development (Abell, 2007; Supovitz & Turner, 2000) as well as low elementary teacher content knowledge and self-efficacy in CS (e.g., Mason & Rich, 2019; Yadav et al., 2016; Rich et al., 2021). This model included a summer academy, academic-year follow-up, and microcredentialing, all aligned with state CS standards.

After one year, the study found statistically significant improvements in teacher CS content knowledge and self-efficacy, suggesting the ARCS model effectively equipped teachers, especially in rural schools, with the knowledge and confidence needed to implement CS instruction (e.g., Li et al., 2025; Rich et al., 2021).

Furthermore, significant gains in student CS content knowledge among treatment schools support the mechanism discussed in the literature that better-prepared teachers lead to stronger student outcomes (Boyd et al., 2009). Although the increase in student interest was not statistically significant, the small positive effect indicates potential for impact with longer-term or more intensive exposure, consistent with research on STEM persistence (Dickerson et al., 2014; Ireland et al., 2018).

Furthermore, the integration of CS into core content and use of inclusive, culturally responsive teaching key strategies emphasized in the literature (Brown-Jeffy & Cooper, 2011; Margolis et al., 2008) were central to the ARCS model and may have contributed to the positive outcomes observed. Overall, the study findings support the conclusion that the ARCS model is well-aligned with the mechanisms needed to improve equity, teacher capacity, and student learning in elementary CS education, particularly in underserved rural settings.

For the ARCS program, next steps include investigating the full two-year model to understand how the second year, especially microcredentialing and ongoing support, adds to teacher growth, student learning, and lasting use of CS instruction in classrooms.

For the field more broadly, these results highlight how important it is to invest in long-term, supportive professional development that helps elementary teachers build confidence, knowledge, and skills in teaching CS. Future work should explore how models like ARCS can be adapted and scaled in different school settings including rural, urban, and under-resourced schools to ensure all students have access to high-quality CS education.

Fidelity of Implementation Study

Fidelity Measurement

Fidelity was measured for 2 key components of the ARCS intervention in year 1 (attendance in coaches academy and academic year follow-up) and 1 component in year 2 (microcredential completion).

For year 1 components, the sample-level fidelity threshold was defined for teacher attendance at the 5-day Code VA Coaches Academy and academic year follow-up engagement. For the coaches academy and academic year follow-up, teacher attendance was aggregated to the school level, and the sample-level fidelity threshold was based on school attendance rates.

Yae 2 fidelity was also measured for micro-credentialing activities and was based on completion of microcredentials.

Fidelity was measured for all schools implementing the intervention each year.

Table 9. Scoring that Defines Adequate Implementation of Each Key Component in a Program Logic Model

Indicator	Unit of measurement	Indicator Scoring at Unit Level	Indicator Scoring at School Level	Indicator Scoring at Sample Level
Key Component 1. Year 1 CODE VA K-5 Coaches Academy (Summer PD)				
Attendance at 5-day Coaches Academy	Teacher	Teacher Level: Attended at least 4 days get a 1	School level: 75% of teachers attended at least 4 days	Program Level: 75 % of schools with a 1
Key Component 2. Year 1 Academic Year Follow Up (Learning Bytes, Coach, Interactions, Tech Sessions, NIC)				
Follow up engagement	Teacher	Teacher Level: Completed at least 3 academic year activities	School level: 75% of teachers get a 1	Program Level: 75 % of schools with a 1
Key Component 3. Year 2 Microcredential Completion				
MC	Teacher	Teacher level: Completes 4 or 5 MC= 1		Program Level: 75 % of teachers with a 1

Fidelity Findings

Cohort 1 (2021-2023)

- Code VA Coaches Academy: 89% of schools met the school-level threshold. There was high fidelity at teacher, school, and program levels for this component.
- Academic Year Follow-Up: 3.6% of teachers engaged in at least 3 academic years follow up activities (i.e., met threshold). There was low fidelity at the teacher, school, and program levels for this component. Program level fidelity not met.
- Microcredentials: 10 of 28 eligible teachers (45.5%) completed at least 4 microcredentials, which did not meet the program level fidelity threshold for this component of ARCS.

Cohort 2 (2023-2025)

- Code VA Coaches Academy: 92% of schools met the school-level threshold. There was high fidelity at teacher, school, and program levels for this component.
- Academic Year Follow-Up – 89% of teachers engaged in at least 3 academic years follow up activities (i.e., met threshold). 96% of schools met the threshold. There was high fidelity at the teacher, school and program levels for this component.
- Microcredentials: 16 of 43 eligible teachers (37.2%) completed at least 4 microcredentials, which did not meet the program level fidelity threshold for this component of ARCS.

Overall Fidelity (both Cohorts of Teachers)

- Code VA Coaches Academy: Across 2 cohorts, 91% of schools achieved fidelity (30/33 schools)
- Academic Year Follow Up: Across 2 cohorts, 73% of schools achieved fidelity (24/33 schools)
- Microcredentials: Across 2 cohorts, 37% of teachers achieved fidelity (26/71)

Table 10. Findings on Fidelity of Implementation by Component in Multiple Years									
Key Components, Number of Indicators, Units, and Threshold				Year 1 Results *(2021-23 School Year)*			Year 2 Results *(2023-25 School Year)*		
Key Component	Total # of Measurable Indicators	Unit of Implementation	Sample-Level Threshold for Fidelity of Implementation	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold	Number of Units in Which Component was Implemented	Number of Units in Which Fidelity of Component was Measured	Achieved Fidelity Score and Whether Program Met Sample-Level Threshold
1.CodeVA Coaches Academy	1 (Attendance)	Teacher (Unit), School Level, and Program Level of Schools	Teacher level: Attended ≥ 4 days=1 School level: ≥ 75% of teachers at least attended 4 days Program level: ≥ 75% of schools met teacher level threshold	28 teachers 9 schools 1 program	28 teachers 9 schools 1 program	Teacher level fidelity: 100% (28/28). School level fidelity: 89% (8/9) Program level fidelity achieved	59 teachers 24 schools 1 program	59 teachers 24 schools 1 program	Teacher level fidelity: 93% (55/59) School level fidelity 92% (22/24) Program level fidelity achieved

2.Follow Up (Learning Bytes, Coach Interactions, tech sessions, NIC)	1 (engagement)	Teacher (Unit), School, and Program	Teacher Level: Completed at least 3 activities = 1 School Level: ≥ 75% of teachers completed at least 3 activities Program Level: ≥ (75% of schools met teacher level threshold	28 Teachers 9 schools 1 program	28 teachers 9 schools 1 program	Teacher-level fidelity: 3.6% (1/28). School-level fidelity: 11.1% (1/9). Program-level fidelity not achieved	59 teachers 24 schools 1 program	55 teachers 24 schools 1 program	Teacher-level fidelity: 89% (48/54). School-level fidelity: 96% (23/24). Program-level fidelity achieved
3.Microcredentials	1 (completion)	Teacher (Unit), Program level	Teacher Level: Teacher completes 4 or 5 microcredentials Program Level: ≥ (75% of teachers eligible for microcredentials with a 1	28 Teachers 1 program	28 teachers 1 program	Teacher-level fidelity: 45.5% (10/28) Program-level fidelity not achieved	43 teachers 1 program	43 teachers 1 program	Teacher-level fidelity: 37.2% (16/43) Program-level fidelity not achieved