# STREAMin$^3$ Curriculum Model
## Evaluation Report

**Virginia Commonwealth University**
Dr. Kevin Sutherland
Dr. Bryce D. McLeod
Dr. Michael Broda
Dr. Rosalie Corona
Dr. Erica Ross
Navneet Kaur
Alisandra Macias
Meghan Reichel
Jeanie Stouffer

# EXECUTIVE SUMMARY

This report summarizes an independent evaluation of the STREAMin³ curriculum model using quantitative and qualitative data collected during the 2022-2023 school year. In 2022, program developers at the University of Virginia (UVA) contacted Virginia Commonwealth University (VCU) to conduct an independent evaluation of the STREAMin³ curriculum during the first year of its statewide availability. A team at VCU led by Dr. Kevin Sutherland, Dr. Bryce D. McLeod and others evaluated the implementation and effectiveness of the STREAMin³ curriculum model. This report describes findings from the following aims:

- **Aim 1**: To examine the effectiveness of STREAMin³ on child, teacher, and classroom outcomes during its first year of implementation.
- **Aim 2**: Explore whether teachers who receive STREAMin³ training and coaching will have higher self-efficacy and lower burnout over the year.
- **Aim 3**: Determine whether classrooms where STREAMin³ is implemented via training and coaching have higher classroom quality than classrooms that did not receive training and coaching.
- **Aim 4**: To examine the association between the dosage of training/coaching and delivery of STREAMin³ with improved classroom outcomes.
- **Aim 5**: To examine factors associated with the dosage of delivery of STREAMin³.
- **Aim 6**: Explore how teacher, classroom, and program characteristics influence the delivery of STREAMin³.

The evaluation team used the Consolidated Framework for Implementation Research (CFIR; Damschroeder et al., 2022) to inform the design and conduct of the evaluation. This report provides an overview of the evaluation. We first describe the evaluation design, including data sources aligned with the study aims. We then discuss participant selection and data collection, including a description of the follow-up interviews with early childhood education teachers. Next, we discuss the data analytic approach used to address each aim, followed by results presented by the aim. We conclude by discussing program considerations based on our findings, limitations of our approach, and final thoughts regarding implications.

Broadly, findings suggest that teachers who implemented STREAMin³ and received more coaching had improved Classroom Assessment Scoring System (CLASS; Pinta et al., 2008) scores compared to those who received less coaching. Further, teachers who received more coaching improved their self-efficacy and could implement more curriculum elements than the teachers who received less coaching. Teachers reported on the strengths of the STREAMin³ approach, particularly the quality of the training and coaching, but also noted challenges, such as difficulties finding time to meet with coaches and implementing the curriculum in mixed-aged classrooms.

## Glossary of Terms/Acronyms

| | |
|---|---|
| **BAU** | "Business As Usual"; name for the sample of teachers that were not using STREAMin³ nor received any training or coaching in STREAMin³ |
| **CFIR** | Consolidated Framework for Implementation Research (CFIR; Damschroder et al., 2022) |
| **EBP** | Evidence-based program |
| **ECE** | Early childhood education |
| **LinkB5** | Data system that includes data on all early childhood education (ECE) classrooms in the Commonwealth |
| **Sample 1** | Matched group that includes teachers who received any STREAMin³ support (Tier 2 and Tier 3) compared to BAU |
| **Sample 2** | Matched group that includes Tier 2 compared to Tier 3 teachers |
| **STREAMin³** | Curriculum being evaluated |
| **Tier 2** | Sample of teachers that received group coaching supports |
| **Tier 3** | Sample of teachers that received individual coaching supports |
| **VKRP** | Virginia Kindergarten Readiness Program |

# INTRODUCTION

## Evaluation Aims

The overarching aim of this evaluation was to examine the effect of STREAMin³ on child outcomes (e.g., literacy, math, science, and socio-behavioral outcomes) associated with school readiness. Additional aims focused on the impact of the curriculum on key teacher outcomes and examined questions about factors associated with STREAMin³ implementation. Specifically, the aims of this evaluation were:

- **Aim 1**: To examine the effectiveness of STREAMin³ on child, teacher, and classroom outcomes during its first year of implementation.
- **Aim 2**: Explore whether teachers who receive STREAMin³ training and coaching will have higher self-efficacy and lower burnout over the course of the year.
- **Aim 3**: Determine whether classrooms, where STREAMin³ is implemented via training and coaching, will have higher classroom quality than classrooms that did not receive training and coaching.
- **Aim 4**: To examine the association between the dosage of training/coaching and delivery of STREAMin³ with improved classroom outcomes.
- **Aim 5**: To examine factors associated with the dosage of delivery of STREAMin³.
- **Aim 6**: Explore how teacher, classroom, and program characteristics influence the delivery of STREAMin³.

Importantly, it should be noted that this evaluation was conducted as an effectiveness study and implemented under typical conditions. Given the design of the evaluation readers should be cautious about making causal inferences. As such, throughout the report, we will attempt to add context to our findings that are relevant for interpreting the results.

## Guiding Framework

The Consolidated Framework for Implementation Research (CFIR; Damschroder et al., 2022) was used to design and conduct the evaluation. Given the effect that schools, and in our case, early childhood settings, have on evidence-based program delivery (Sutherland et al., 2008), it is important to better understand how these contexts affect implementation. The CFIR uses an ecological framework to help researchers determine how factors in a particular context (i.e., early childhood programs, classrooms) influence the implementation of EBP and is comprised of five major domains: (a) intervention characteristics; (b) outer setting (e.g., the broader economic, social or political context within which a school sits); (c) inner setting (e.g., characteristics of the school); (d) characteristics of those delivering the EBP (e.g., teacher experience and training); and (e) the implementation process (e.g., training, coaching). The CFIR encourages researchers to think about how the characteristics of an EBP interact with inner and outer settings and the characteristics of those delivering the EBP to best adapt the EBP to fit the specific needs of the children, teachers, and program context. McLeod et al. (2020) point out that the CFIR is well-situated to inform research in education since implementation research in school and early childhood contexts is still in its infancy, and the CFIR is designed to promote theory development and verification.

## Sections of the Report

This report provides an overview of the evaluation. We first describe the evaluation design, including data sources aligned with the study aims. We then discuss participant selection and data collection, including a description of the follow-up interviews with ECE teachers. Next, we discuss the data analytic approach used to address each aim, followed by results presented by the aim. We conclude by discussing program considerations based on our findings, limitations of our approach, and final thoughts regarding implications.

# Evaluation Design

## Evaluation Aims

**Table 1**

*Study aims and associated data sources*

| Study Aims | Data Sources | | | | |
|---|---|---|---|---|---|
| | VCU Surveys | LinkB5 | VKRP | UVA Surveys | VCU Interviews |
| *To examine the effectiveness of STREAMin³ on child, teacher, and classroom outcomes.* | | **X** | **X** | **X** | |
| *To explore whether teachers who receive STREAMin³ training and coaching will have higher levels of self-efficacy and lower levels of burnout across the course of the year.* | **X** | | | | |
| *To determine whether classrooms where STREAMin³ is implemented via training and coaching will have higher classroom quality than classrooms that did not receive training and coaching.* | **X** | **X** | | **X** | |
| *To examine the association between the dosage of training/coaching and delivery of STREAMin³ with improved classroom outcomes.* | **X** | | | | |
| *To examine factors associated with the dosage of delivery of STREAMin³.* | **X** | | | | |
| *Explore the mechanisms through which teacher, classroom, and program characteristics influence the delivery of STREAMin³.* | | | | | **X** |

## Participant Selection

Our target population of teachers included 300 early learning and care classrooms participating in the individual coaching condition of STREAMin³ identified by the UVA STREAMin³ team. We began the recruitment process by emailing all eligible participants, describing the purpose of the study and the surveys that would be sent out in the coming months. Participants were then sent

a link via Qualtrics to complete a Pre-Assessment survey. Initial recruitment efforts began in November 2022 and continued through January 2023 resulting in a final sample of 167 completed surveys. Table 2 below describes the teacher demographics of this sample and Tables 11 and 12 in the appendix give further details about classroom and student demographics.

**Table 2**
*Teacher Demographics for Tier 3 Coaching Sample*

| | VCU Survey Sample (*n* =167) | Qualitative Interview Sample (*n* = 18) |
|---|---|---|
| **Gender** | | |
| Female | 162 (97.0%) | 17 (94.4%) |
| Male | 2 (1.2%) | 0 (0.0%) |
| Non-binary/third gender | 1 (0.6%) | 0 (0.0%) |
| Prefer not to answer | 2 (1.2%) | 1 (5.6%) |
| **Race** | | |
| White | 77 (46.1%) | 8 (44.4%) |
| African American/Black | 70 (40.9%) | 9 (50.0%) |
| Asian/Pacific Islander | 5 (3.0%) | 0 (0.0%) |
| Native American/American Indian | 1 (0.6%) | 0 (0.0%) |
| Other | 10 (6.0%) | 0 (0.0%) |
| Prefer not to Answer | 4 (2.4%) | 1 (5.6%) |
| **Hispanic/Latinx** | | |
| No | 149 (89.2%) | 16(88.9%) |
| Yes | 18 (10.8%) | 2 (11.1%) |
| **Teacher/Age** | | |
| 18-25 | 26 (15.6%) | 0 (0.0%) |
| 26-35 | 42 (25.1%) | 6 (33.3%) |
| 36-45 | 46 (27.5%) | 4 (22.2%) |
| 46-55 | 31 (18.6%) | 4 (22.2%) |
| Over 55 | 22 (13.2%) | 4 (22.2%) |
| **Highest Level of Education** | | |
| High School Diploma | 69 (41.3%) | 5 (27.8%) |
| Associate's Degree | 47 (28.1%) | 4 (22.2%) |
| Bachelor's Degree | 25 (15.0%) | 5 (27.8%) |
| Master's Degree | 6 (3.6%) | 0 (0.0%) |
| Education Specialist | 2 (1.2%) | 0 (0.0%) |
| Other | 18 (10.8%) | 4 (22.2%) |
| **Licensed** | | |
| No | 108 (64.7%) | 10 (55.6%) |
| Yes | 59 (35.3%) | 8 (44.4%) |
| **Years of Teaching Experience** | | |

| | | |
|---|---|---|
| 5 or less | 62 (37.2%) | 6 (33.3%) |
| 6-10 | 38 (22.8%) | 4 (22.2%) |
| 11-15 | 25 (15.0%) | 2 (11.1%) |
| 16-20 | 23 (13.8%) | 4 (22.2%) |
| 21-25 | 6 (3.6%) | 1 (5.6%) |
| 26-30 | 9 (5.4%) | 1 (5.6%) |
| More than 30 | 3 (1.8%) | 0 (0.0%) |
| Missing | 1 (0.6%) | 0 (0.0%) |
| **Years of Teaching Pre-K Experience** | | |
| 5 or less | 90 (54.0%) | 7 (28.8%) |
| 6-10 | 30 (18.0%) | 3 (16.7%) |
| 11-15 | 17 (10.2%) | 3 (16.7%) |
| 16-20 | 15 (9.0%) | 4 (22.2%) |
| 21-25 | 7 (4.2%) | 0 (0.0%) |
| 26-30 | 5 (3.0%) | 1 (5.6%) |
| More than 30 | 3 (1.8%) | 0 (0.0%) |
| Missing | 0 (0.0%) | 1 (5.6%) |

Probe 1 was sent out in January 2023. The initial recruitment of this survey resulted in 56 completed surveys. The team followed up with participants via email or phone for Probe 1 at four different time points between January and February, with each round of follow-up improving survey completion: 43 (follow-up 1), 8 (follow-up 2), 3 (follow-up 3), and 5 (follow up 4) resulting in 115 completed surveys. Probe 2 was administered in February and March. The initial recruitment of this survey resulted in 37 completed surveys. Four separate follow-ups were made via emails and phone calls. Each round of follow-ups improved survey completion: 46 (follow-up 1), 12 (follow-up 2), 0 (follow-up 3), and 1 (follow-up 4). Probe 3 was sent out in March and April. The initial recruitment of this survey resulted in 12 completed surveys. Two separate follow-ups were made via emails and phone calls. Each round of follow-ups improved survey completion: 47 (follow-up 1) and 12 (follow-up 2). The final survey was sent out in April resulting in 45 completed surveys. Three separate follow-ups were made via emails and phone calls. Each round of follow-ups improved survey completion: 7 (follow-up 1), 4 (follow-up 2), and 4 (follow-up 3). This resulted in a final sample of 60 teachers who completed the Pre-Assessment, Probes, and Post-Assessment. Table 10 in the appendix gives further information about our recruitment efforts.

The evaluation team also received secondary data from the University of Virginia's STREAMin³ team. This data allowed the team to create matched samples to different groups of varying STREAMin³ support to address our research aims. Descriptions of these samples can be found below in Table 3 with a more detailed description found in the Data Collection and Analysis section.

**Table 3**

*Participant Names and Descriptions*

| Sample Names | Descriptions |
|---|---|
| *Sample 1: VCU Survey Participants* | Participants who received Tier 3 STREAMin³ Individual Coaching support and participated in the VCU survey. |
| *Sample 2: Business As Usual (BAU) and Tier 2 and Tier 3 STREAMin³ Support.* | The BAU sample included a state sample of teachers who did not receive any support or enrolled in STREAMin³. The Tier 2 and Tier 3 sample included participants who received either Tier 2 (Group Coaching) or Tier 3 (Individual Coaching) STREAMin³ Coaching. The BAU teachers were used as a comparison sample to identify differences in outcomes for teachers who received any STREAMin³ support. |
| *Sample 3: Tier 2 STREAMin³ Support and Tier 3 STREAMIn³ Support* | This sample only includes participants who received either Tier 2 (Group Coaching) or Tier 3 (Individual Coaching) STREAMin³ Coaching. For this sample, the Tier 2 group was used as a comparison sample to identify differences in outcomes for teachers depending on the amount of coaching support. The Tier 3 teachers are also some of the same participants as the VCU Survey and Interview samples. We were unable to link the data to connect these datasets due to not having access to identifying information. |
| *Sample 4: VCU Participant Interviews* | Teachers who participated in the interviews following selection based on analyzing teachers' AIM/IAM/FIM scores. |

## Data Collection and Analysis

There were five sources of data for this evaluation: (1) VCU Surveys for the Tier 3 Individual Coaching Group, (2) VCU Interviews, (3) LinkB5 data, (4) VKRP data, and (5) UVA Survey data for all other participants receiving STREAMin³ support.

### Study Design and Data Collection Procedures

The VCU evaluation team collected two primary sources of data: survey data and teacher interviews conducted via Zoom. The team used a case study design with a sequential explanatory approach (Creswell, 2014). This involves collecting and analyzing the quantitative data first, followed by qualitative interviews to address the study's aims.

### Data Collection Procedures

Pre- and post-assessments were conducted along with three probes. To evaluate the effectiveness and implementation of STREAMin³, the pre-assessment included STREAMin³ Teacher Demographic and Classroom Composition questions, the Intervention Usability Scale (IUS), the Acceptability of Intervention Measure (AIM), Intervention Appropriateness Measure (IAM), and Feasibility of Intervention Measure (FIM), STREAMin³ specific implementation

questions as well as coaching dosage, the Teacher Sense of Self- Efficacy Scale (TSES), and the Maslach Burnout Inventory (MBI). The probes were intended to assess teachers' perspectives on the feasibility, acceptability, and appropriateness of STREAMin³ along with coaching dosage, across the course of the school year.

**Figure 1.** ***Data collection timeline***


**Measures**

*Intervention Usability Scale* (IUS; Lyon et al., 2021) is a 10-item teacher-report scale that assesses the usability of EBPs. A recent evaluation of the adapted IUS found good internal consistency (α=.83) and the ability to differentiate among psychosocial interventions (Lyon et al., 2021).

*Acceptability, Feasibility, and Appropriateness*. Teachers rated the feasibility, acceptability, and appropriateness of STREAMin³ using the 4-item *Feasibility of Intervention Measure* (FIM), 4-item *Intervention Appropriateness Measure* (IAM), and 4-item *Acceptability of Intervention Measure* (AIM) at each time point. Each measure has demonstrated internal consistency, validity, and responsiveness to change (Weiner et al., 2017).

Coaching dosage. Teachers reported on how often and for how much time they met with their coaches.

*Teachers' Sense of Self-Efficacy Scale* (TSES; Tschannen-Moran & Hoy, 2001) includes three sub-scales that measure teacher efficacy related to student engagement, instructional strategies, and classroom management. Scale reliabilities range from .87 to .94.

*Maslach Burnout Inventory* (MBI; Maslach et al., 1997) assesses teacher burnout and consists of three sub-scales: Depersonalization, Emotional Exhaustion, and Personal Accomplishment, which have demonstrated high internal consistency ranging from .71 to .90.

## Interview Procedures

From the sample of 60 teachers who completed all five waves of data collection, 34 teachers were invited for semi-structured interviews based on their scores on the acceptability, feasibility, and appropriateness (AIM/IAM/FIM) survey. Recruitment for interviews started at the end of April via an initial email detailing information related to the interviews and compensation. If the teacher did not reply to the initial email, follow-up calls and emails were sent. Of the teachers contacted, 18 were interviewed via Zoom by three research staff (see Table 2 for teacher demographics).

To identify teachers for the interviews, they were grouped based on their AIM/IAM/FIM scores to create a sample of teachers with differing perspectives on the curriculum. Four groups were identified based on whether teacher scores on the AIM/FIM/IAM were high or low over time (high and low perceptions, respectively) or if their scores increased or decreased (positive and

negative slopes, respectively). Of the 18 teachers who completed the interview, 4 teachers were in the negative slope group, 6 in the positive slope group, 4 in the high perceptions group, 3 in the low perceptions group, and 1 teacher who consented to be interviewed did not fit into one of the four groups.

We created the interview questions using the CFIR (Damschroder et al., 2022). Teachers were asked about their experiences with the coaching, training they received, and their implementation of the curriculum in their classrooms. Interviews were conducted via Zoom, lasting from 26 to 66 minutes.

## Secondary Data

The evaluation team also received secondary data from the University of Virginia's STREAMin³ team. This included state-level LinkB5 data, the Virginia Kindergarten Readiness Program (VKRP), and data UVA collected from all STREAMin³ participants across the state. LinkB5 data includes measures of the quality of teaching and learning in every publicly funded ECE classroom in the Commonwealth. VKRP data includes a coordinated set of assessments that measure children's skills in early literacy (VLP), mathematics (Early Mathematics Assessment System; EMAS), self-regulation (Child Behavior Rating Scale; CBRS), and social skills (CBRS). This data allowed the team to use propensity score matching (see below) to create matched control groups for treatment groups to account for covariates and allow the groups to have similar characteristics to reduce bias. Following the matching process, we used the matched samples of secondary data to (1) examine the effectiveness of STREAMin³ on the child, teacher, and classroom outcomes and (2) determine whether classrooms receiving STREAMin³ training and coaching will have higher classroom quality than classrooms that did not receive training and coaching.

## Analyses

To best address our research questions, the team used propensity score matching to match participants in the treatment groups with the most similar participants in the comparison group (Zhao et al., 2021). The team created two matched groups– (1) a treatment group that included teachers who had received either STREAMin³ Tier 2 or Tier 3 coaching support, and (2) a control group, referred to as our business as usual (BAU) group, that included teachers that did not receive any STREAMin³ support. Matching these two groups generated Sample 2.

A second matching process allowed us to examine the differences between teachers who received different levels of coaching support. For these analyses, the treatment group included teachers who received Tier 3 coaching support, while the control group included teachers who received Tier 2 coaching support. Matching these groups generated Sample 3.

The team used the LinkB5, VKRP, and UVA Survey data to create our matched samples. The UVA survey data included a variable indicating the type of coaching the participant received from UVA. This variable allowed the team to create our two matched groups. We then used the MatchIt function in R to match our two samples. For Sample 2, each teacher that received any STREAMin³ coaching support (Tier 2 or Tier 3) was matched with a BAU teacher. For Sample 3, each teacher who received Tier 3 individual coaching was matched with a teacher who received Tier 2 group coaching support. The following variables were used to create the matched samples for both groups: Does the site offer state or federal food programs to kids

(food_program_yes), whether transportation is available at the site (transportation_no), whether there are specific eligibility requirements child must meet to enroll in program (eligibility_req_no), whether the site has a waitlist (child_on_waitlist_yes), whether Title 1 is a revenue source (title1_no), whether Early Childhood Special Education (ECSE or IDEA) is a revenue source (ecse_idea_no), whether Headstart is a revenue source (headstart_no), whether Early Headstart is a revenue source (early_headstart_no), whether VPI is a revenue source (vpi_no), whether the Child Care Subsidy Program (VA CCSP) is a revenue source (vaccsp_yes), whether Child Care Access Means Parents in School (CCAMPIS) is a revenue source (ccampis_no), and whether the site has a VA Quality rating (va_quality_rating_earned_yes). Tables 13 and 14 in the appendix include more details about the propensity score matching scores.

# Results

## *Aim 1: To examine the effectiveness of STREAMin³ on child, teacher, and classroom outcomes.*

Using a quasi-experimental design, the team examined whether children exposed to STREAMin³ have higher literacy, mathematics, science, and social/emotional/behavioral (SEB) skills than children in the BAU condition. The research team used Sample 2 to examine differences between all teachers that received STREAMin³ coaching support for the 2022-2023 year and the BAU sample. We then further examined if there were differences based on the type of coaching support by using Sample 3. The team explored descriptive statistics, t-tests, and linear regression to control for Fall scores and compare means between the two groups.

Table 4 provides 2023 Spring raw scores for child behavior (CBRS subscales), math (EMAS sub-scales) and reading (the PALs Assessment subscales) for the matched sample that includes BAU and all teachers that received Tier 2 and Tier 3 STREAMin³ coaching support (Sample 2). We also include a *p*-value for the Spring score difference after controlling for Fall scores. ***Patterning was significantly higher for the STREAMin³ Coaching sample compared to those that did not receive STREAMin³ support.*** Table 15 in the appendix provides descriptive statistics of all subscales analyzed.

**Table 4**

*VKRP Scores: Difference between receiving STREAMin³ Coaching and BAU*

| | BAU (*n* = 412-1,090) | STREAMin³ Coaching (*n* = 120-262) | BAU (*controlling for Fall)* | STREAMin³ Coaching (*controlling for Fall)* | *p*-value | Cohen's *d* | Lower CI | Upper CI |
|---|---|---|---|---|---|---|---|---|
| **VKRP: CBRS Self-Regulation Score: Spring** | 3.69 (0.560) | 3.72 (0.455) | 0.40 | 0.34 | 0.32 | 0.054 | -0.094 | 0.035 |
| **VKRP: CBRS Social Skills Score: Spring** | 4.01 (0.489) | 4.00 (0.418) | 0.20 | 0.18 | 0.48 | 0.019 | -0.049 | 0.068 |
| **VKRP: Patterning Subdomain Score: Spring** | 2.66 (0.984) | 2.88 (0.962) | 0.29 | 0.46 | 0.001*** | 0.220 | -0.347 | -0.084 |
| **VKRP: Numeracy Subdomain Score: Spring** | 9.99 (3.26) | 9.89 (3.15) | 2.79 | 2.22 | 0.01** | 0.373 | -0.332 | 0.529 |
| **PALS: Beginning Sounds Matching (0-10): Spring** | 3.29 (2.15) | 3.64 (2.27) | 1.05 | 0.671 | 0.55 | 0.162 | -0.813 | 0.106 |
| **PALS: Letter Names (0-52): Spring** | 34.5 (10.3) | 34.0 (10.9) | 17.40 | 16.400 | 0.5 | 0.048 | -0.959 | 1.959 |
| **PALS: Letter Sounds (0-25): Spring** | 12.5 (5.98) | 12.6 (6.22) | 8.53 | 8.230 | 0.15 | 0.018 | -0.947 | 0.725 |

*Note.* Regression analyses were run for each variable to control for Fall scores to find p-values. ** p < .01, *** p<.001.

Table 5 provides the same data but for the matched sample (Sample 3) that only includes STREAMin³ teachers - Tier 2 Group Coaching as the control and Tier 3 individual coaching as the intervention group. We explored these differences further through comparing VKRP scores for the different levels of coaching. There were no significant differences between the Tier 2 and Tier 3 coaching groups, although after controlling for Fall scores, ***the Tier 3 Individual Coaching group had higher mean scores for the PALS letter sounds and letter names.***

**Table 5**

*VKRP Scores: Difference between Tier 2 and Tier 3 Coaching Groups*

| | Tier 2 Group Coaching (*n* = 36-40) | Tier 3 Direct Coaching (*n* = 25-26) | Tier 2 Group Coaching (*controlling for Fall)* | Tier 3 Direct Coaching (*controlling for Fall)* | *p*-value | Cohen's *d* | Lower CI | Upper CI |
|---|---|---|---|---|---|---|---|---|
| **VKRP: CBRS Self-Regulation Score: Spring** | 3.72 (0.44) | 3.60 (0.47) | 0.21 | 0.06 | 0.28 | 0.28 | -0.11 | 0.36 |
| **VKRP: CBRS Social Skills Score: Spring** | 3.87 (0.40) | 3.79 (0.40) | 0.09 | 0.10 | 0.41 | 0.21 | -0.12 | 0.28 |
| **VKRP: Patterning Subdomain Score: Spring** | 3.53 (1.03) | 3.34 (1.06) | 0.43 | 0.43 | 0.49 | 0.18 | -0.35 | 0.72 |
| **VKRP: Numeracy Subdomain Score: Spring** | 9.79 (3.28) | 9.26 (3.27) | 0.53 | 0.04 | 0.52 | 0.16 | -1.12 | 2.18 |
| **PALS: Beginning Sounds Matching (0-10): Spring** | 4.17 (2.45) | 4.01 (2.00) | 0.49 | 0.32 | 0.80 | 0.07 | -1.07 | 1.38 |
| **PALS: Letter Names (0-52): Spring** | 28.1 (11.6) | 29.2 (12.3) | 9.38 | 9.60 | 0.71 | 0.10 | -7.41 | 5.06 |
| **PALS: Letter Sounds (0-25): Spring** | 9.74 (6.30) | 11.2 (7.05) | 4.50 | 4.73 | 0.42 | 0.22 | -4.93 | 2.07 |

*Note.* Regression analyses were run for each variable to control for Fall scores to find p-values.

## *Aim 2: Explore whether teachers who receive STREAMin³ training and coaching will have higher levels of self-efficacy and lower levels of burnout across the course of the year.*

The research team used the VCU data (Sample 1: Tier 3 coaching group) to compare mean scores in self-efficacy and burnout between the fall and spring. Teachers in the Tier 3 coaching group who acknowledged having no coaching support the month prior to the survey were used to compare differences in self-efficacy and burnout with the population of teachers who had received coaching support the month prior. Coaching dosage was also considered and descriptively explored across self-efficacy and burnout for the teachers who indicated they had received coaching within the month prior to the spring survey.

Tables 6 and 7 and Figures 2 and 3 include data from the VCU surveys that only included the STREAMin³ Tier 3 Coaching group. Maslach's Burnout Scale (MBI) scores and Teacher's Sense of Self Efficacy (TSES) Scores were compared between the Fall of 2022 and Spring of 2023 and were also compared across dosage of coaching support. ***Teachers reported higher emotional exhaustion, higher depersonalization, and lower personal accomplishment in the Spring compared to the Fall.*** It is important also to consider that there are confounding variables that could impact this such as poor wages, work-family conflict, and parental demands

(De Los Santos et al., 2023). ***But despite higher burnout scores, teachers reported a higher sense of self-efficacy across all subscales.***

**Table 6**

*MBI and TSES Scores of Tier 3 Teachers in the Fall and Spring*

| | Fall 2022 Scores (*n* = 165-167) | Spring 2023 Scores (*n* = 60) |
|---|---|---|
| **Burnout** | | |
| Emotional Exhaustion | 22.30 | 24.80 |
| Personal Accomplishment | 48.80 | 47.50 |
| Depersonalization | 7.30 | 7.67 |
| **Sense of Self-Efficacy** | | |
| Classroom Management | 7.36 | 7.42 |
| Instructional Strategies | 7.27 | 7.38 |
| Student Engagement | 7.58 | 7.62 |

We further explored burnout and self-efficacy by comparing the Spring scores of teachers who indicated they had met with their coach in the past month to those who had not. ***Teachers who met with their coach in the last month (n = 44) reported lower emotional exhaustion and depersonalization, and higher personal accomplishment than those who had not (n = 16). Teachers who reported meeting with their coach in the last month also had statistically significantly higher self-efficacy scores than those who did not.***

**Table 7**

*MBI and TSES Scores of Tier 3 Teachers Who Met with Their Coach the Month Before the End of the Year and Those Who Did Not*

| | Met With Coach (*n* = 44) | Did Not Meet with Coach (*n* = 16) | *p*-value | Cohen's *d* | Lower CI | Upper CI |
|---|---|---|---|---|---|---|
| **Burnout** | | | | | | |
| Emotional | 23.2 | 29.3 | 0.12 | 0.48 | -13.96 | 1.70 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Exhaustion | | | | | | |
| Personal Accomplishment | 47.6 | 47.0 | 0.78 | 0.08 | -3.94 | 5.17 |
| Depersonalization | 7.05 | 9.38 | 0.08 | 0.66 | -4.95 | 0.29 |
| **Sense of Self-Efficacy** | | | | | | |
| Classroom Management | 7.64 | 6.80 | 0.02* | 0.67 | 0.12 | 1.57 |
| Instructional Strategies | 7.58 | 6.84 | 0.05* | 0.64 | 0.01 | 1.46 |
| Student Engagement | 7.84 | 7.02 | 0.02* | 0.71 | 0.17 | 1.48 |

We descriptively looked at the sample of teachers who reported in the Spring having met with their coaches the month before, and examined burnout and self-efficacy based on the dosage of coaching they reported receiving (Figures 2 and 3). Due to small sample size, these results should be interpreted with caution. The four groups had relatively similar results, with the teachers who reported receiving less than 30 minutes of coaching having slightly higher emotional exhaustion and lower perception of instructional strategies compared to teachers who received over 30 minutes of coaching.

**Figure 2.** ***Distribution of MBI Across the Amount of Coaching Received***

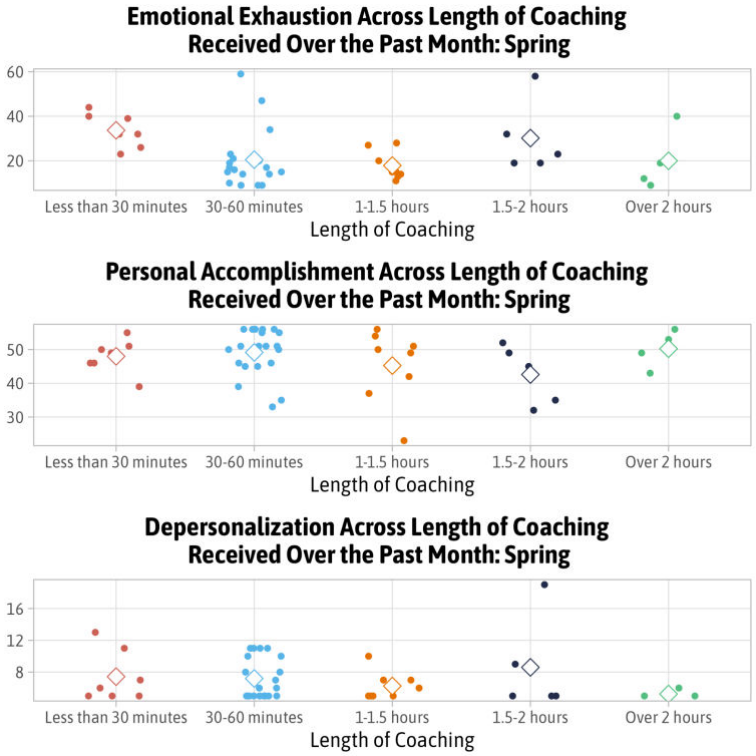
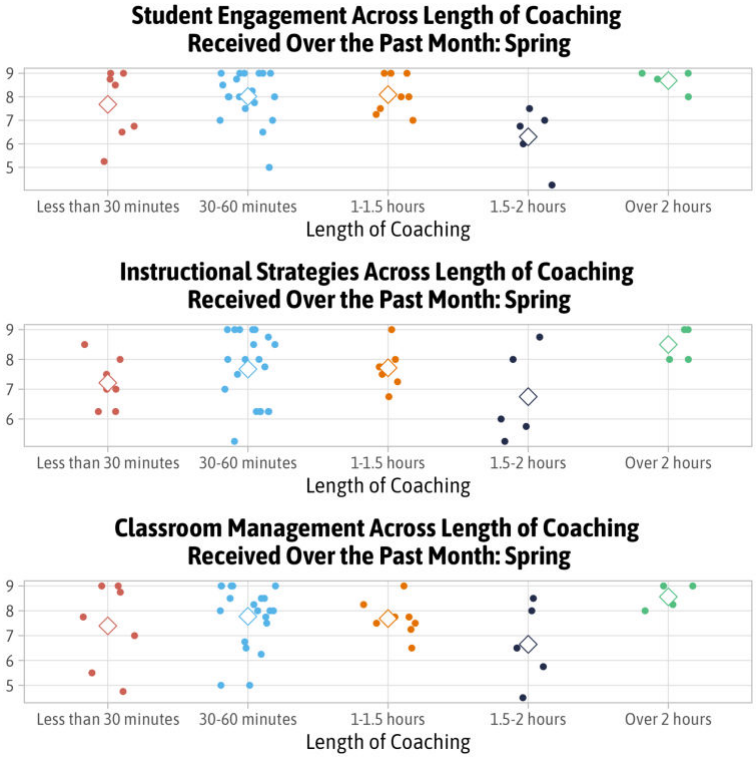



**Figure 3.** ***Distribution of TSES Across the Amount of Coaching Received***



## *Aim 3: Determine whether classrooms where STREAMin³ is implemented via training and coaching will have higher classroom quality than classrooms that did not receive training and coaching.*

Sample 2 was used to compare means between the groups to determine differences in classroom quality using the Classroom Assessment Scoring System (CLASS; Pinta et al., 2008) data found in LinkB5. The research team used Sample 2 to examine differences between all teachers that received STREAMin³ coaching support for the 2022-2023 year and the BAU sample. We also examined if there were differences in CLASS scores based on the type of coaching support using Sample 3. The team explored descriptive statistics, t-tests, and linear regression to control for fall scores and compare means between the two groups.

Data from Table 8 provides 2023 Spring raw scores for subscales of the CLASS for the matched sample that includes BAU and all teachers that received Tier 2 and Tier 3 STREAMin³ coaching support (Sample 2). Although there were not statistically significant differences between the BAU classrooms and those that received any STREAMin³ coaching support, when we controlled for Fall scores by subtracting the mean Fall score from the Spring score, we found that ***those classrooms that had STREAMin³ coaching support had a higher adjusted mean for instructional support and engaged support for learning.***

**Table 8**

*CLASS Scores: Difference between* receiving STREAMin³ Coaching and BAU

| | BAU (N = 9,971) | STREAMin³ Coaching (N = 2,176) | *p*-value | Cohen's *d* | Lower CI | Upper CI |
|---|---|---|---|---|---|---|
| **CLASS Responsive Caregiving Domain: Spring** | 5.121 | 5.371 | 0.52 | 0.247 | -0.420 | -0.080 |
| **CLASS Emotional Behavior Support Domain: Spring** | 5.694 | 5.862 | 0.44 | 0.204 | -0.260 | -0.075 |
| **CLASS Instructional Support Domain: Spring** | 3.450 | 3.614 | 0.67 | 0.137 | -0.274 | -0.053 |
| **CLASS Emotional Support Domain: Spring** | 6.102 | 6.184 | 0.83 | 0.122 | -0.142 | -0.021 |
| **CLASS Engaged Support for Learning Domain: Spring** | 3.608 | 3.889 | 0.099 | 0.233 | -0.427 | -0.134 |
| **CLASS Classroom Organization Domain: Spring** | 5.539 | 5.626 | 0.54 | 0.087 | -0.175 | 0.001 |

*Note.* Regression analyses were run for each variable to control for Fall scores.

Table 9 provides the same data but for the matched sample that only includes STREAMin³ teachers (Sample 3). After controlling for Fall scores, ***teachers who received individual coaching (Tier 3) had significantly higher average scores for emotional support and classroom organization than teachers who received group coaching (Tier 2).*** While not significantly different, the Tier 3 group also had higher mean scores across all CLASS domains.

**Table 9**

*CLASS Scores: Differences between Tier 2 and Tier 3 Coaching*

| | Tier 2 Coaching (N = 951) | Tier 3 Coaching (N = 300) | *p*-value | Cohen's *d* | Lower CI | Upper CI |
|---|---|---|---|---|---|---|
| **CLASS Responsive Caregiving Domain: Spring** | 5.23 | 5.51 | 0.17 | 0.305 | -0.63 | 0.07 |
| **CLASS Emotional Behavior Support Domain: Spring** | 5.74 | 5.97 | 0.28 | 0.308 | -0.41 | -0.05 |
| **CLASS Engaged Support for Learning Domain: Spring** | 3.80 | 3.99 | 0.73 | 0.163 | -0.48 | 0.09 |
| **CLASS Emotional Support Domain: Spring** | 5.97 | 6.22 | 0.006** | 0.364 | -0.43 | -0.07 |
| **CLASS Classroom Organization Domain: Spring** | 5.30 | 5.65 | 0.04* | 0.381 | -0.59 | -0.11 |
| **CLASS Instructional Support Domain: Spring** | 3.25 | 3.45 | 0.81 | 0.17 | -0.50 | 0.10 |

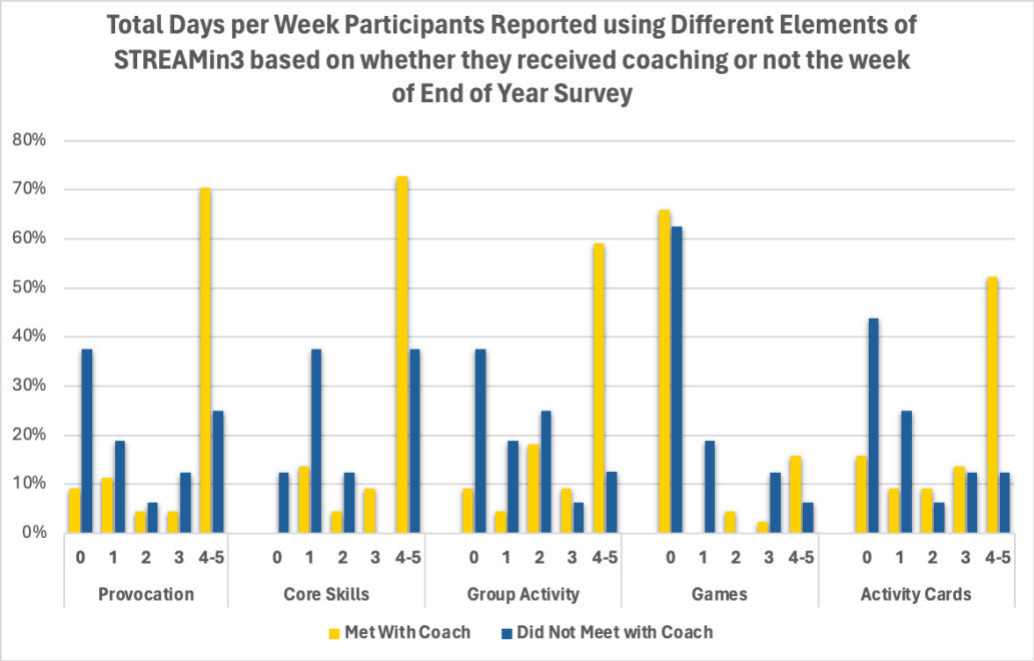
*Note.* Regression analyses were run for each variable to control for Fall scores. *p < .05, **p < .01.

## *Aim 4: To examine the association between dosage of training/coaching and delivery of STREAMin³ with improved classroom outcomes.*

The VCU team collected data capturing the dosage of coaching and the frequency of use of STREAMin³. These data were used to examine whether there are descriptive differences between the dosage of coaching and whether teachers are using the program with the intended frequency.

The team explored descriptive differences between the dosage of coaching and whether teachers were using the program with the frequency intended to be used within the Tier 3 individual coaching sample through survey data captured by VCU. Figure 4 shows the total days per week participants reported using elements of the STREAMin³ curriculum based on whether they had met with their coach the month prior. It should be noted that the curriculum elements are aimed to be incorporated 4-5 days a week. ***Across all five curriculum elements, those teachers who met with their coach had a higher percentage of use within the desired adherence than those who did not meet with their coach.*** Table 16 in the Appendix has a more detailed response frequency and percentages table.

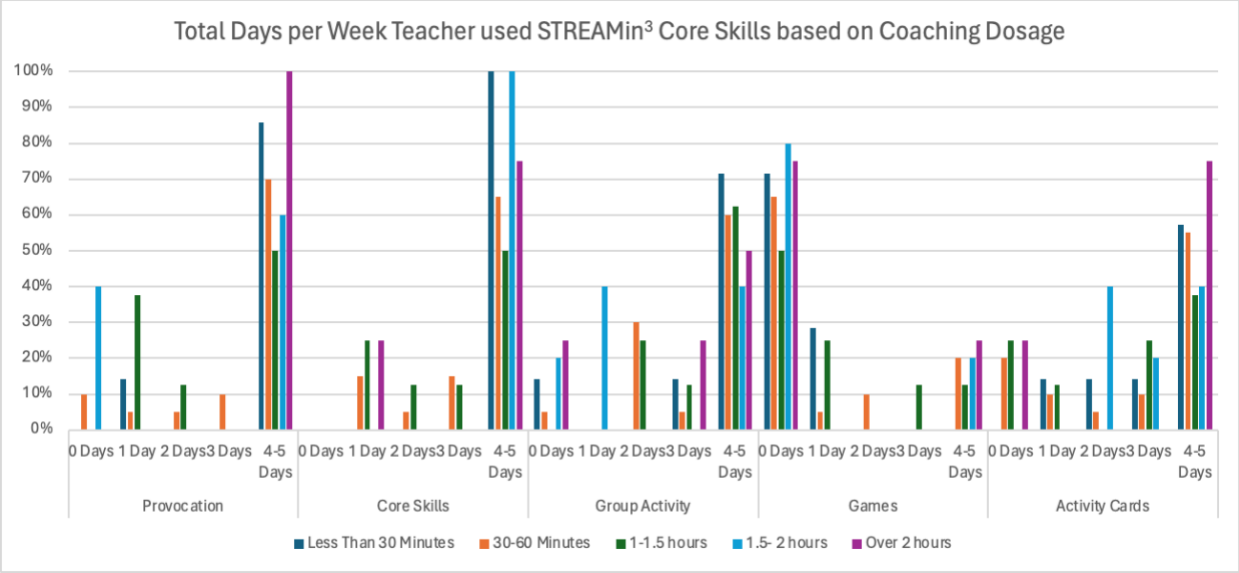
**Figure 4.** ***Total Days per Week Participants Reported Using Elements of STREAMin³***




*Note.* See Appendix, Table 17 for a more detailed table of the responses.

We further explored the total days per week participants reported using different elements of STREAMin³ based on the length of coaching they received. Despite differences in coaching dosage, most participants indicated using the different curriculum elements within the desired adherence.

**Figure 5.** ***Total Days per Week Participants Reported Using Core Skills of STREAMin³***



## *Aim 5: To examine factors associated with the dosage of delivery of STREAMin³.*

VCU survey data and data captured through LinkB5 were used to examine teacher (experience, training), classroom level, and program (type) associations with dosage of delivery of STREAMin³.

The team used VCU survey data to examine teacher (experience, training) and classroom level associations with dosage of delivery of STREAMin³ (See Appendix, Tables 18-22). We captured experience through years of teaching and training through teaching licensure and education level. We used whether they met with their coach or current position as a classroom factor. For this analysis, we define delivery dosage as how many days a week teachers implement the STREAMin³ curriculum. ***Meeting with their coach was significantly and positively associated with implementing the curriculum for almost every element.*** Being a Pre-K teacher was significantly and positively associated with the activity cards and the games. This aligns with the program aims as both of these elements are present in the Pre-K curriculum compared to the infant curriculum.

## *Aim 6: Explore the mechanisms through which teacher, classroom, and program characteristics influence delivery of STREAMin³.*

Responses to the qualitative questions were transcribed using an AI service (TEMI). Transcripts were checked for accuracy by the research staff. Transcripts were uploaded into Dedoose, a qualitative software program. Three members of the team conducted a six-step thematic analysis of the data (Braun & Clark, 2006). (1) Data familiarization. First, the team familiarized themselves with the data by reading and re-reading the transcripts in detail to look for emerging codes (e.g., processes, actions, assumptions, consequences). Codes also included metaphors, repetitions across interviewees, and shifts in content that indicated relevant themes (Strauss & Corbin, 1990). While codes mainly originated from participants' transcripts (inductive), creating codes also included prior knowledge of implementation science (deductive). The three research team members met to discuss our impressions and ideas for potential codes. (2) Code creation. Next, we generated the initial codes and defined them. A codebook was created to aid with coder reliability, and example quotes were provided for each code. Two team members coded each transcript, and the three members met to discuss discrepancies (i.e., consensus coding). The two coders also wrote memos about the coding to maintain transparency and develop an audit trail. (3) Collation of codes. Codes were then collated and reviewed by the coding team. (4) Theme development. After coding was completed, themes were developed. Themes are constructs that researchers identify before, during, and after data collection and can be generated from myriad pathways, including the data collected, literature reviews, and researchers' experiences with the data (Polkinghorne, 2005; Ryan & Bernard, 2000). Four themes emerged from the data (see below): Supports, Barriers/Challenges, Effectiveness, and Recommendations. (5) Review, revise, and finalize themes. The team continued to meet to discuss, review, and finalize the themes. (6) Finally, a report was written summarizing the qualitative findings.

Four themes emerged from the data: (1) Supports, (2) Barriers, (3) Effectiveness, and (4) Recommendations. The Supports and Barriers themes each had sub-themes. Themes were interrelated. For instance, teachers described factors supporting their use of the curriculum (e.g., helpful coaches and trainers), and these factors were linked with recommendations they provided. Similarly, recommendations stemmed from the barriers teachers described in

implementing the curriculum.

**Theme 1: Supports**

Factors that supported the use of the curriculum in the classrooms included helpful coaches and trainers, the onboarding/training (e.g., improving the understandability of the curriculum, adaptations are acceptable), being provided with the curriculum and resources, and having supportive colleagues and organization management.

**Helpful coaches and trainers.** Teachers shared many positive comments about their coaches and the trainers. Teachers enjoyed the in-person training and noted that the trainers were knowledgeable, supportive, and helpful. They said that trainers helped them better understand the curriculum and how to implement it in their classrooms. Teachers' communication with coaches took many forms, including in-person meetings, Zoom, and phone calls. Teachers talked about positive experiences with their coaches, including their support, reassurance, and availability. Teachers also appreciated that coaches helped them generate ideas to implement the curriculum in their specific classrooms. Sample quotes from teachers included:

- *They did a great job…they made it [the curriculum] so approachable, so simple and practical…I really enjoyed the in-person trainings.*
- *The kind words…she knows her job, so she was real good at coaching. You know, when I thought I was wrong or I wasn't doing the right, but she was, she reassured was 100%.*
- *Ms. Coach is my coach and she's very helpful and very patient because…I'm not going at a fast pace, a steady pace. But she's very helpful with her suggestions on ways to, um, intertwine the two age levels…and, um, she's constantly, you know, giving me information about different resources that will help me do that.*

**Improved understandability.** Teachers shared that their training before implementing the curriculum and the deeper dive modules improved their understanding. The training also increased teachers' comfort in using the curriculum in their classrooms. This was especially true for teachers who found the curriculum overwhelming. Sample quotes included:

- *When she went through the ACT activity cards…she went basically page for page, you know, and, um, she explained, you know, how to use the activity card, and I thought that was, that was good cause that helped me a whole lot.*
- *It [deeper dive videos] broke it down in a better way to understand it.*
- *So the trainings…I've been to the in-person trainings. Those are phenomenal…how they were able to put it together so easily. It's different from like reading and watching the training videos versus being in person and then you're all in one room, like, oh, I get it. And then seeing what everybody else's best practices are and like, okay, I like that I could apply that…that could help me improve this*

**Adaptations are okay.** Teachers also appreciated the messages from the coaches and their trainers that it is okay to modify the curriculum to fit their classroom. Teachers talked about their modifications, including using their materials when implementing the curriculum with children of different ages and English-limited students. Sample quotes included:

- *What I took away from most of the training was that this curriculum is made to be modified and I feel like in the written part of the curriculum, that's not easily understood.*
- *Um, yes because it's, we are a Spanish immersion preschool program and we also use ASL…so I had to incorporate or do a lot of things differently, like move a lot of the activities outside or use other books as well that weren't in, um, as part of the curriculum because we, we read a lot in Spanish.*

- *Um, [coaching] actually made a bigger difference…better understanding it and how, how to apply it and how to make, um, changes that need to be changed for certain kids…like I had this one kid who wasn't, couldn't use paint, so I had to improvise with that and she kind of helped, gave me different ideas on how to improvise for him. Um, cause he has eczema so he couldn't really touch the paint because of all the chemicals in it. And so she kind of helped me, um, find different ways for him to do it so that he wasn't left out so he could participate too.*

**Building Teachers' Toolbox.** Teachers noted that the link between their teaching skills and the skills in the curriculum made it easier to implement the curriculum in their classes. Teachers also found being provided with the curriculum and materials (e.g., books, flashcards, Tucker the Turtle) for implementing the curriculum helpful. Sample quotes included:

- *I already have like a nature-based part of my program, so the curriculum does include like planting seeds like in herb garden and then transplanting some things outside. So that was helpful to have it already like, uh, as part of the, of the curriculum cause we were doing that already anyways.*
- *Um, there for a while I struggled with like the language modeling but, um, I'm working with a coach now to try to get better at that cause I think I scored kind of low on that on my observation…Sometimes I want you to say things I'm not supposed to say, but I have to stop myself and go, no, you know…you can't do that cause they're not gonna understand. So that's been a work in progress and I think I'm doing better on that.*
- *Oh yeah. Like, um, normally, you know, before STREAMin[3], we probably just would've gone on a leaf hunt. I wouldn't have, you know, asked questions like, "oh, is this one big or small" or you know, "is this one crunchy?" Or you know… that the leaf has fallen off the tree type of thing. Um, so it definitely was helpful.*

**Peer/Organizational Support.** The support of other teachers or leadership in an organization further facilitated teacher's use of the curriculum in their classrooms. Sample quotes included:

- *Well, they were pretty much open-minded to use this as it was the new thing that we all had to use. So everyone, we all kind of had to accept it and figure it out together.*
- *Um, well, all the employees, we kind of help each other. Like if there's more than one twos teacher, you know, one of us will, you know, we kind of help each other out and if we have any questions, some of 'em understand it better than others so we can go to the ones that understand it better. They help us through it or give us ideas on how we can do, um, like bird, I think it was like the bean, we had to plant beans one day, so, you know, they said, "well hey, this is an idea for twos, this might be easier for twos." So, we all kind of just worked together and it made, it made it easier.*
- *Um, we start off doing different activities…each teacher had a different day than we did and we worked together, figure out when, like when we had our learning time, when we'll do it and tell the children what we're doing today. That way they better understand things.*

**Theme 2: Barriers**

Teachers were generally positive about the coaches, training, and curriculum. However, they did share some challenges they encountered when implementing the curriculum in their classrooms. These implementation challenges/barriers included child (e.g., mixed-age classroom) and curriculum-related (e.g., coaching, training, materials/resources) factors.

**Child Factors.** A few teachers who were interviewed ($n$=4) had children of different age groups in the same classroom (e.g., infants and toddlers). Teachers with mixed-age classrooms found it difficult to implement the curriculum because children were at different developmental stages.

Further, teachers who primarily worked with younger children (e.g., infants) reflected that the curriculum was not fully developed for the younger children. For example:

- *The most challenging part about the whole thing…is, um, having mixed ages. So it's, it can be somewhat challenging sometimes to make sure you're covering one part of the day with the preschooler and then you got your toddler part that they're getting their part in. The curriculum is great, it's just the time management when you have mixed ages.*
- *A lot of it could have been implemented more, but where my kids were so young, there was, there's only just the, like a little section that I could actually use…But, uh, like as far as just going by the book curriculum, um, I think that was probably the most challenging part, was just the fact that my, my children were on the lower age.*
- *If this program is good for preschoolers… but for babies, I don't think it's working.*

**Coaching Challenges.** When teachers shared their challenges with the coaching, it was not about the coaches themselves but more about the coaching process. Specifically, teachers noted that it was hard to find time for coaching due to their busy schedule. Teachers also found communication through emails, the amount of emails, and having a coach far away as challenging. One teacher noted that she preferred to talk to her coach, but she recognized that the coach might be with another teacher, so she opted to send an email. One teacher said that a challenging part was having her coach be far away. Sample quotes included:

- *Um, I think the scheduling, but that's just gonna be for whatever. Um, I mean it wasn't really challenging but she was good at being flexible at scheduling it, but just in general just taking the time to schedule another thing.*
- *I've received so many emails that I'm kind of confused with who the emails are from really. Cause I received emails from my coach. I received emails from my inspector I received, and, and, and they are almost the same emails coming from different, um, companies…You know, I'm just getting same emails, but it's the same emails, but it's from different companies. And, and it's be, it's confusing. So, my email be so full that I say, "oh, I'm not gonna look through this today. I'm just not gonna look through."*
- *I think it's our, our timing as far as, sometime I might have had another training or something that overlapped, so I think it's just the time management of us getting together. Um, sometimes we have, I have crossed trainings or something that have, may have missed it, so, but she's been very accommodating. You know, by rescheduling it and yes.*

**Trainings.** Teachers discussed feeling overwhelmed with the training and that the training was confusing. Some teachers noted that the virtual platform did not work well and reported difficulty understanding what the trainers were talking about and identifying what in the curriculum they were referring to. For example:

- *I think at the beginning, cause it's so many books and so many things…I don't recall that the training had like the first steps, like where to look before we even get into the week one. Maybe you did, but I, because I was looking at week one is where I think I would've started week, and not being at week one was a challenge for me.*
- *Um, we did the training, guess it was the beginning of the year where it was, um, like through a Zoom type thing and our whole department was together. Um, and to be honest, none of us liked it…And it was just…not very organized, um, because there are different levels of STREAMin$^3$, like someone would ask where it was in the book and it was different in all the books, the pages and all that. So it was just very confusing. (laugh) is, I guess the best way to put it.*
- *Um, it was a lot of information. I feel like, um, the trainers sometimes it was hard to follow cause it seemed like they were reading from a script and would kind of go back*

*and forth between the different, um, books and materials. And so sometimes it was just hard to keep up with, with like which book we were supposed to be looking at the time. But, um, overall it was a lot of, a lot of good information.*

**Understandability of the Curriculum.** Some teachers found the curriculum to be overwhelming in length and understandability. However, as noted above, teachers gained a lot from and appreciated their coaches, who helped them understand and implement the curriculum in their classrooms. Sample quotes include:

- *Implementing this curriculum is the hardest thing because even if you can understand, um, how wordy it is and how the theories make sense implementing it is not clearly stated.*
- *Uh, the wording…I like it short and simple. and they would take a sentence, maybe a paragraph, and it would stretch it so long that I kinda like forget the thought. Cause...it was just too long. So, um, that was mainly the main thing, just like I said, that reading part.*
- *When I first, uh, received the STREAMin[3], I did not understand it. Truly, I didn't. And some of the terms words that was used, I still didn't understand it because it was a lot of reading and I wanted to see, I wanted to see it, um, everyday language. I wanted to see it in good examples and I just couldn't see it when I first got it. And I just put the book the box down and I wouldn't read it.*

**Theme 3: Program Effectiveness**

Teachers shared that the curriculum improved their teaching and that the children enjoyed the activities. Teachers also shared that the skills taught helped prepare the children for future schooling and furthered their development of social-emotional skills. The theme of program effectiveness is linked with the Supports: Building Teacher's Toolbox sub-theme. Specifically, teachers talked about how the curriculum helped them integrate new teaching skills into their teaching toolbox and helped prepare them for working with diverse groups of children (e.g., children with disabilities). For example:

- *I like how it helps you like ask more open-ended questions with the kids, um, because that was something that I struggled with and it, you know, it gives you examples of how to ask those type of questions with the activities.*
- *I wouldn't change anything about the curriculum. I think it's very good. And, um, one of my friends, she works in the daycare and they're using the same curriculum. I said, "oh, we're doing that too," you know…So I thought that was cool…I think she has like four to five year-olds and she said how that has helped prepare the children, you know, going into kindergarten. So I think that it's a good curriculum…it helps if the child has a disability or something, you can spot it before they get to school. So that's what I like about it because if that child is having a problem, then, you know, they can catch it and they can pass it on to the parents. So that's been really good.*
- *The things under like regulate that has helped a whole lot. Um, and then relate, they've, they've helped me better relate to a child's like feelings and trying to understand them more, even if they can't speak, it kind of helps you to figure out ways to try to try to figure out what the child is saying and so that's really helped me a whole lot.*

**Theme 4: Recommendations**

Not surprisingly, the teachers' recommendations were connected to the themes of Supports and Barriers. Below are specific recommendations for coaching and training.

**Coaching.** Teachers recommended (a) continued coaching and access to coaches, (b) increased time availability of coaching sessions so that it does not interfere with classroom time and weekends, and (c) scaffolding in teaching the curriculum. Sample quotes included:

- *That was a lot of material to receive all at one time. And so having the mentor gradually say, "Hey, have you done this?" Starting from the beginning and then working up, not even going into the curriculum, which, um, that was, um, awesome with our, uh, mentor just following up, "Hey, do you have any questions? I'm always here and always available. What was your opinion on this that you tried?"*
- *Um, I think if we would've had more coaching in the beginning…To really set the foundation, it would've helped us later on.*
- *I'm a curriculum person, but for those that have never seen a curriculum, I believe that, uh, mentor is needed a lot longer than the time that was given. Highly needed.*

**Training.** Recommendations for training/onboarding included (a) smaller groups during the training sessions, which would allow more time to review the curriculum and deeper dive videos, and (b) separate training for the separate age modules. In addition, a few participants expressed a desire for in-person training, while others enjoyed the flexibility of the virtual options. Sample quotes include:

- *Um, like I said, to me the, the best thing that they could do is to break it out by the age groups and do separate trainings…That to me makes the most sense.*
- *I'm just saying, um, if it was up to me, I would have diving deep mandatory.*
- *So if they could do more workshops on that [relate], that would be awesome too.*

**Materials Provided.** Teachers appreciated receiving the materials and resources for implementing the curriculum. Their recommendations centered on (a) continuing to provide teachers with these resources, including board games, (b) packaging the curriculum and materials together, and (c) providing alternative resources for different ages and situations (e.g., providing resources in Spanish for immersion programs). One teacher noted that she could not complete a lesson because she forgot that she needed live worms. This type of comment suggests that having lessons not tied to teacher-sourced materials may improve implementation. Teachers also noted that the curriculum was too "wordy" and hard to understand. Their recommendations were to simplify the curriculum and include more examples and alternatives. Another teacher suggested adding a "parent portal" so that learning could continue at home and to support parent-teacher conferences. For example:

- *Well definitely having the physical games and I think that's part of the problem that like just having a, a little book to tell you the games I really just go and grab my own physical board games or, you know, activities to implement.*
- *The only thing I think that can be improved, it's probably, I know it's new, so the process of letting the providers know that materials are coming. Because when I got the two cases of books, I had no idea where they were coming from and I was like, I can't do nothing with these books. And I was going, I was going to give them away but I received a phone call from someone from STREAMin[3] and she said, "by the way, did you get your books?" And I, and I said, "what books?" She said, "the books from Barnes and Noble." I was like, "oh." So that's where these two cases of books came from. So, I think by being new it could just been a loss. Yeah. I think that was like, where, where is the stuff coming from? And then the, the curriculum came after that.*
- *So the curriculum as a whole, the way that it was written, (laugh), um, just the way that it was, the way that it was written was the main thing. If we could, if it was written more simply, simplified, I think it would be more successful of a curriculum.*

# Evaluation Summary

The purpose of this report was to summarize an independent evaluation of the STREAMin³ curriculum model using quantitative and qualitative data collected during the 2022-2023 school year. The evaluation team used quantitative and qualitative data to examine the curriculum's effectiveness and factors associated with the implementation of STREAMin³ using the CFIR (Damschroder et al., 2022) to design and conduct the evaluation. Importantly, and to provide context, this was an independent evaluation of the curriculum that occurred under naturally occurring circumstances; that is, from a translational perspective, this is an effectiveness study with little to no experimental control provided. Thus, treatment effects and findings should be interpreted within this context.

Below, we will first discuss the evaluation findings, focusing on teacher and child outcomes. We will then discuss some of the findings associated with varying amounts of coaching and factors associated with implementation, followed by a discussion of the themes found in the qualitative data. Notably, we hope to integrate our findings so that this report is helpful to the developers of STREAMin³ and the Commonwealth as it makes determinations about curricular use moving forward. We will finish our discussion with the limitations of our evaluation.

Importantly, there were positive effects on child outcomes in classrooms that used STREAMin³ compared to those that didn't. Specifically, there were significant effects, with small effect sizes (*d* of .22)for Patterning, respectively. Of note, there were also mean differences favoring the STREAMin³ group for Self-Regulation and Beginning Sounds; the remaining child outcomes were essentially the same for the two groups. Relatedly, there were also mean differences, favoring the STREAMin³ group, on all CLASS domains. While these were not statistically significant, they were consistent.

The findings here are enhanced if we examine the differences between the two coaching groups, Tier 2 versus Tier 3. First, there were no statistically significant differences between the two groups, with mean differences favoring the Tier 3 group for PALS Letter Names and Letter Sounds. While these findings should be viewed cautiously due to a small sample, it does suggest that the benefits of more intensive coaching may not be expressed in child outcomes. However, these findings are different if we examine the CLASS scores between the two coaching groups. Interestingly, CLASS scores for all domains were higher in classrooms where teachers received Tier 3 support versus those who received Tier 2 coaching, and two of the domains (Emotional Support and Classroom Organization) were significantly higher in Tier 3 classrooms, with effects sizes of .36 and .38, respectively. This seems important and consistent with the universal, comprehensive approach of the STREAMin³ curriculum. That is, teachers who receive the most training and coaching appear more able to enhance their classroom quality.

These findings are also consistent with teachers' sense of self-efficacy as measured by the TSES. Teachers who received Tier 3 coaching supports increased their self-efficacy across the school year for each of the three domains (Classroom Management, Instructional Strategies and Student Engagement), and our findings also suggest that teachers who received more coaching (represented by teachers who received coaching in the last month of school versus those did not) had even higher self-efficacy, with effect sizes ranging from .64 to .71. In addition, while teachers who received Tier 3 supports throughout the year did not show improvements in the burnout domains as assessed by the MBI, when we examined teachers who received more coaching (again represented by teachers who received coaching in the last month of school

versus those did not), we see reductions in Emotional Exhaustion and Depersonalization, although these findings were not statistically significant (likely due to a lack of statistical power given the small sample size; effect sizes were .48 and .66, respectively). In addition, teachers who met with their coaches more often had a higher percentage of use of the STREAMin³ curriculum elements, and for almost every element of the curriculum, meeting with their coach was significantly and positively associated with implementing the curriculum. These findings suggest that teachers who receive more coaching feel more efficacious in managing their classrooms and supporting children's learning. They were less exhausted and were thus able to create higher-quality learning contexts. Child outcomes may take more time to be seen in these environments, but overall, the learning context appears improved.

The data from the qualitative interviews largely supported the quantitative findings. We sampled teachers, all of whom received Tier 3 coaching support, who had varying perspectives on the curriculum as well as varying levels of implementation. As mentioned earlier in this report, four themes emerged from the qualitative data: Supports, Barriers, Effectiveness, and Recommendations. Teachers noted the coaches' helpfulness and the training's quality, although time constraints were mentioned as barriers to receiving coaching (i.e., finding time). In addition, some teachers noted that the groundwork for implementation could be laid earlier to make the uptake of training easier.

Other strengths of the STREAMin³ curriculum noted by teachers included the 'built-in' acceptance of adaptations. This seemed particularly important for teachers to meet the needs of English-language learners and children with disabilities, and teachers appreciated that they could modify the curriculum for specific learner needs. Teachers also noted how STREAMin³ and the coaching helped support their skills development. These findings support the abovementioned results about teachers' self-efficacy- the more coaching they received, the more efficacious they felt in meeting children's needs in the classroom.

Of course, teachers also had specific recommendations for improving the curriculum and discussed barriers to implementation. One of the most significant barriers mentioned came from teachers with mixed-aged classrooms that included toddlers and infants, with teachers reporting more challenges implementing the infant curriculum.

## Limitations and Final Thoughts

There are several limitations to keep in mind as readers interpret these findings. First, some challenges were associated with accessing the data for the full sample, and missing data made the matching approach difficult. This resulted in a limited sample size for some aims, reducing statistical power. A lack of statistical power limits our ability to detect certain findings associated with the more indicated Tier 3 sample. Finally, while we could sample teachers who received Tier 3 coaching based on their perspectives of STREAMin³ and their implementation, we did not have time to fully integrate the qualitative and quantitative findings in a true mixed method approach. This concern is mitigated due to the richness of the qualitative data and the variety of teachers whom we sampled and interviewed.

Finally, it is important to return to the context of this evaluation, represented by the effectiveness design and the CFIR, which highlights multiple factors that influence the implementation of EBPs such as STREAMin³. It is imperative to point out that this evaluation and the data collected therein occurred during the teachers' first year of implementation of a new, comprehensive curriculum. It is likely (based on data from our own work) that teachers will

become even more comfortable with implementing STREAMin³ in subsequent years. Thus, the promising results seen in this evaluation are likely to grow, and we would encourage investments in coaching as this does seem to be a critical implementation support for ECE teachers in the Commonwealth.

# References


Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, *13*(1), 3-21.

Creswell, J. W. (2014*). Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (4th ed.). Thousand Oaks, CA: Sage.

Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O., & Lowery, J. (2022). The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation science*, *17*(1), 75. https://doi.org/10.1186/s13012-022-01245-0

De Los Santos, R., Borchardt PsyD, J. N., Yousey, B., Dickson, S., Aloise, S., Butler, M., & Banker Ed D, D. (2023). A narrative review of preschool teacher burnout. *Modern Psychological Studies*, *29*(2), 1.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1997). *Maslach burnout inventory*. Scarecrow Education.

McLeod, B. D., Kunemund, R., Nemer, S. L., & Lyon, A. R. (2020). Leveraging implementation science and practice to support the delivery of evidence-based practices in services for youth with emotional and behavioral disorders. In *Handbook of research on emotional and behavioral disorders* (pp. 417-432). Routledge.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*. Paul H. Brookes Publishing Co.

Polkinghorne, D. E. (2005). Language and meaning: Data collection in qualitative research. *Journal of counseling psychology*, *52*(2), 137–145. https://doi.org/10.1037/0022-0167.52.2.137.

Ryan, G. W., & Bernard, H. R. (2000). Techniques to identify themes in qualitative data.

Sutherland, K. S., Lewis-Palmer, T., Stichter, J., & Morgan, P. L. (2008). Examining the influence of teacher behavior and classroom context on the behavioral and academic outcomes for students with emotional or behavioral disorders. *The Journal of Special Education*, *41*(4), 223-233. https://doi.org/10.1177/0022466907310372

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, *17*(7), 783-805.Lyon, A. R., Pullmann, M. D., Jacobson, J., Osterhage, K., Al Achkar, M., Renn, B. N., ... & Areán, P. A. (2021). Assessing the usability of complex psychosocial interventions: The Intervention Usability Scale. *Implementation Research and Practice*, *2*, 2633489520987828.

Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation science: IS*, *12*(1), 108. https://doi.org/10.1186/s13012-017-0635-3

Zhao, B., & Bilen, H. (2021, July). Dataset condensation with differentiable siamese augmentation. In *International Conference on Machine Learning* (pp. 12674-12685). PMLR.

# Appendix

**Table 10**

*Participant Recruitment: Number of Completed Surveys and Dates of Recruitment*

| | Pre-Assessment | | Probe 1 | | Probe 2 | | Probe 3 | | Post- Assessment | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Date | # | Date | # | Date | # | Date | # | Date | # |
| Initial Recruitment | 11/7/22 *(Cohort 1)* 1/12/23 *(Cohort 2)* | 110 | 1/09/23 *(Cohort 1)* 2/2/23 *(Cohort 2)* | 56 | 2/6/23 *(Cohort 1)* 2/16/23 *(Cohort 2)* | 37 | 3/14/23 *(Cohort 1 & 2)* | 12 | 4/3/23 *(Cohort 1 & 2)* | 45 |
| Follow-up 1 | 11/14/22 *(Cohort 1)* 1/19/23 *(Cohort 2)* | 13 | 1/16/23 *(Cohort 1)* 2/5/23 *(Cohort 2)* | 43 | 2/13/23 *(Cohort 1)* 2/23/23 *(Cohort 2)* | 46 | 3/21/23 *(Cohort 1 & 2)* | 47 | 4/10/23 *(Cohort 1 & 2)* | 7 |
| Follow-up 2 | 11/21/22 *(Cohort 1)* 1/26/23 *(Cohort 2)* | 11 | 1/24/23 *(Cohort 1)* 2/9/23 *(Cohort 2)* | 8 | 2/20/23 *(Cohort 1)* | 12 | 3/28/23 *(Cohort 1 & 2)* | 12 | 4/17/23 *(Cohort 1 & 2)* | 4 |
| Follow-up 3 | 11/29/22 *(Cohort 1)* | 3 | 1/31/23 *(Cohort 1)* | 3 | 2/28/23 *(Cohort 1)* | 0 | | - | 4/25/23 *(Cohort 1 & 2)* | 4 |
| Follow-up 4 | 12/8/22 *(Cohort 1)* | 24 | 1/31/23 *(Cohort 1)* | 5 | 2/28/23 *(Cohort 1)* | 1 | | - | | - |
| Follow-up 5 | 12/15/22 *(Cohort 1)* | 3 | - | - | | - | | - | | - |
| Follow-up 6 | 12/21/22 *(Cohort 1)* | 3 | - | - | | - | | - | | - |

**Table 11**

*Classroom Demographics*

| Variables | *n* (%) |
|---|---|
| **Classroom Composition** | |
| Pre-Kindergarten (Pre-K) | 101 (60.5%) |
| Mixed/Collaborative | 66 (39.5%) |
| **Number of Adults** | |
| 0 | 3 (1.8%) |
| 1 | 40 (24.0%) |
| 2 | 92 (55.1%) |
| 3 | 21 (12.6%) |
| 4 | 4 (2.4%) |
| 5 | 1 (0.6%) |
| Missing | 5 (3.0%) |

**Table 12**

*Child Demographics*

| | *n* (%) |
|---|---|
| **Ages of Children** | |
| Infant | 12 (7.2%) |
| Toddler | 32 (19.2%) |
| Preschool | 69 (41.3%) |
| More than one of these | 54 (32.3%) |
| **Number of Children** | |
| 5 or less | 28 (16.8%) |
| 6-10 | 76 (45.6%) |
| 11-15 | 37 (22.2%) |
| 16-20 | 17 (10.2%) |
| 21-25 | 3 (1.8%) |
| 26-30 | 3 (1.8%) |
| More than 30 | 1 (0.6%) |
| Missing | 2 (1.2%) |
| **Female Children** | |
| 5 or less | 109 (65.4%) |
| 6-10 | 51 (30.6%) |
| 11-15 | 4 (2.4%) |
| 16-20 | 2 (1.2%) |
| More than 20 | 1 (0.6%) |
| **Male Children** | |
| 5 or less | 104 (62.4%) |
| 6-10 | 52 (31.2%) |
| 11-15 | 3 (1.8%) |
| 16-17 | 2 (1.2%) |
| **Children Classified as English Language Learners (ELL)** | |
| 5 or less | 89 (54.5%) |
| 6-10 | 43 (25.8%) |
| 11-15 | 18 (10.8%) |
| 16-20 | 12 (7.2%) |
| More than 20 | 3 (1.8%) |
| **Children with an Individualized Education Plan (IEP)** | |
| 5 or less | 165 (98.9%) |
| 6-10 | 1 (0.6%) |

| | |
|---|---|
| 11-15 | 1 (0.6%) |

**Table 13**

*Propensity Score Matching Results: Business as Usual (Control) and STREAMin[3] Matched Sample (Treatment)*

| Summary of Balance for All Data | | | |
|---|---|---|---|
| | Means Treated | Means Control | Std. Means Dif |
| food_program_yes | 0.205 | 0.147 | 0.144 |
| transportation_no | 0.102 | 0.078 | 0.079 |
| eligibility_req_no | 0.195 | 0.118 | 0.196 |
| child_on_waitlist_yes | 0.204 | 0.145 | 0.146 |
| title1_no | 0.286 | 0.219 | 0.149 |
| ecse_idea_no | 0.251 | 0.185 | 0.152 |
| headstart_no | 0.311 | 0.219 | 0.198 |
| early_headstart_no | 0.195 | 0.118 | 0.196 |
| vpi_no | 0.224 | 0.162 | 0.149 |
| vaccsp_yes | 0.199 | 0.122 | 0.193 |
| ccampis_no | 0.321 | 0.261 | 0.128 |
| va_quality_rating_earned_yes | 0.145 | 0.101 | 0.124 |
| Summary of Balance for Matched Data | | | |
| food_program_yes | 0.205 | 0.205 | 0 |
| transportation_no | 0.098 | 0.098 | 0 |
| eligibility_req_no | 0.194 | 0.194 | 0 |
| child_on_waitlist_yes | 0.203 | 0.203 | 0 |
| title1_no | 0.284 | 0.284 | 0 |
| ecse_idea_no | 0.250 | 0.250 | 0 |
| headstart_no | 0.308 | 0.308 | 0 |
| early_headstart_no | 0.194 | 0.194 | 0 |
| vpi_no | 0.222 | 0.222 | 0 |
| vaccsp_yes | 0.199 | 0.199 | 0 |
| ccampis_no | 0.319 | 0.319 | 0 |
| va_quality_rating_earned_yes | 0.144 | 0.144 | 0 |
| Sample Sizes | Treated | Control | |
| All | 2186 | 17898 | |
| Matched (ESS) | 2176 | 9971 | |
| Matched | 2176 | 16492 | |
| Unmatched | 10 | 1406 | |

**Table 14**

*Propensity Score Matching for Tier 2 Group Coaching (Control) and Tier 3 Individual Coaching (Treatment)*

| Summary of Balance for All Data | | | |
|---|---|---|---|
| | Means Treated | Means Control | Std. Means Dif |
| food_program_yes | 0.187 | 0.210 | -0.061 |
| transportation_no | 0.079 | 0.109 | -0.107 |
| eligibility_req_no | 0.259 | 0.174 | 0.193 |
| child_on_waitlist_yes | 0.209 | 0.201 | 0.019 |
| title1_no | 0.283 | 0.287 | -0.010 |
| ecse_idea_no | 0.283 | 0.240 | 0.095 |
| headstart_no | 0.283 | 0.320 | -0.083 |
| early_headstart_no | 0.259 | 0.174 | 0.193 |
| vpi_no | 0.281 | 0.204 | 0.170 |
| vaccsp_yes | 0.277 | 0.173 | 0.232 |
| ccampis_no | 0.281 | 0.334 | -0.117 |
| va_quality_rating_earned_yes | 0.148 | 0.144 | 0.012 |
| class_type_family_childcare_provider | 0.251 | 0.025 | 0.522 |
| class_type_private_center | 0.749 | 0.611 | 0.317 |
| **Summary of Balance for Matched Data** | | | |
| food_program_yes | 0.177 | 0.177 | 0 |
| transportation_no | 0.071 | 0.071 | 0 |
| eligibility_req_no | 0.242 | 0.242 | 0 |
| child_on_waitlist_yes | 0.207 | 0.207 | 0 |
| title1_no | 0.255 | 0.255 | 0 |
| ecse_idea_no | 0.255 | 0.255 | 0 |
| headstart_no | 0.255 | 0.255 | 0 |
| early_headstart_no | 0.242 | 0.242 | 0 |
| vpi_no | 0.253 | 0.253 | 0 |
| vaccsp_yes | 0.255 | 0.255 | 0 |
| ccampis_no | 0.255 | 0.255 | 0 |
| va_quality_rating_earned_yes | 0.131 | 0.131 | 0 |
| class_type_family_childcare_provider | 0.238 | 0.238 | 0 |
| class_type_private_center | 0.762 | 0.762 | 0 |
| **Sample Sizes** | Treated | Control | |
| All | 541 | 1649 | |
| Matched (ESS) | 521 | 412.8 | |
| Matched | 521 | 951 | |
| Unmatched | 20 | 698 | |

**Table 15.**

*VKRP Sub-scale Scores for CBRS, EMAS, and PALS for the BAU Sample and STREAMin[3]*

| | BAU (N=16492) | STREAMin[3] Support (N=2178) |
|---|---|---|
| **VKRP: CBRS Self-Regulation Score: Spring** | | |
| Mean (SD) | 3.69 (0.560) | 3.72 (0.455) |
| Median [Min, Max] | 3.73 [1.20, 4.99] | 3.73 [2.14, 4.98] |
| Missing | 15402 (93.4%) | 1916 (88.0%) |
| **VKRP: CBRS Social Skills Score: Spring** | | |
| Mean (SD) | 4.01 (0.489) | 4.00 (0.418) |
| Median [Min, Max] | 4.06 [1.86, 5.00] | 4.01 [2.62, 5.00] |
| Missing | 15402 (93.4%) | 1916 (88.0%) |
| **VKRP: Well-Being Composite: Spring** | | |
| Mean (SD) | 4.26 (0.428) | 4.27 (0.408) |
| Median [Min, Max] | 4.31 [2.37, 5.00] | 4.32 [3.00, 4.98] |
| Missing | 15402 (93.4%) | 1916 (88.0%) |
| **VKRP: Geometry Subdomain Score: Spring** | | |
| Mean (SD) | 5.47 (1.32) | 5.41 (1.22) |
| Median [Min, Max] | 5.57 [0, 8.00] | 5.56 [1.40, 8.00] |
| Missing | 15422 (93.5%) | 1917 (88.0%) |
| **VKRP: Patterning Subdomain Score: Spring** | | |
| Mean (SD) | 2.66 (0.984) | 2.88 (0.962) |
| Median [Min, Max] | 2.70 [0, 6.00] | 2.85 [0.500, 6.00] |
| Missing | 15422 (93.5%) | 1917 (88.0%) |
| **VKRP: Numeracy Subdomain Score: Spring** | | |
| Mean (SD) | 9.99 (3.26) | 9.89 (3.15) |

| | | |
|---|---|---|
| Median [Min, Max] | 10.4<br>[0, 18.0] | 10.3<br>[1.40, 16.6] |
| Missing | 15422 (93.5%) | 1917 (88.0%) |
| **VKRP: Computation Subdomain Score: Spring** | | |
| Mean (SD) | 3.10 (1.17) | 3.15 (1.07) |
| Median [Min, Max] | 3.19<br>[0, 5.00] | 3.30<br>[0.375, 5.00] |
| Missing | 15422 (93.5%) | 1917 (88.0%) |
| **VKRP: Spanish Patterning Subdomain Score: Spring** | | |
| Mean (SD) | 4.87 (2.12) | 7.23 (1.19) |
| Median [Min, Max] | 5.00<br>[0, 8.00] | 7.72<br>[5.50, 8.00] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: Well-Being Composite: Spring** | | |
| Mean (SD) | 2.46 (1.64) | 3.03 (0.0625) |
| Median [Min, Max] | 2.50<br>[0, 6.00] | 3.00<br>[3.00, 3.13] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: Spanish Numeracy Subdomain Score: Spring** | | |
| Mean (SD) | 9.49 (4.89) | 14.8 (1.32) |
| Median [Min, Max] | 10.0<br>[0, 17.0] | 15.0<br>[13.0, 16.2] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: Spanish Computation Subdomain Score: Spring** | | |
| Mean (SD) | 2.94 (1.87) | 3.52 (1.47) |
| Median [Min, Max] | 3.50<br>[0, 5.00] | 3.78<br>[1.50, 5.00] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: Spanish EMAS In-Person Total Score: Spring** | | |
| Mean (SD) | 19.8 (9.32) | 28.6 (3.74) |

| | | |
|---|---|---|
| Median [Min, Max] | 20.5<br>[0, 34.0] | 30.2<br>[23.0, 31.0] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: Spanish EMAS In-Person Scaled Score: Spring** | | |
| Mean (SD) | 574 (97.5) | 664 (50.1) |
| Median [Min, Max] | 572<br>[296, 755] | 671<br>[598, 718] |
| Missing | 16451 (99.8%) | 2174 (99.8%) |
| **VKRP: EMAS English Total Score (in-person and remote combined): Spring** | | |
| Mean (SD) | 21.4 (5.85) | 21.4 (5.58) |
| Median [Min, Max] | 22.0<br>[0.00, 35.0] | 21.9<br>[6.00, 33.2] |
| Missing | 15431 (93.6%) | 1917 (88.0%) |
| **VKRP: EMAS English Scaled Score (in-person and remote combined): Spring** | | |
| Mean (SD) | 587 (66.7) | 584 (66.3) |
| Median [Min, Max] | 591<br>[296, 817] | 591<br>[377, 778] |
| Missing | 15431 (93.6%) | 1917 (88.0%) |
| **PALS: Beginning Sounds Matching (0-10): Spring** | | |
| Mean (SD) | 3.29 (2.15) | 3.64 (2.27) |
| Median [Min, Max] | 3.25<br>[0, 10.0] | 3.31<br>[0, 9.67] |
| Missing | 16080 (97.5%) | 2058 (94.5%) |
| **PALS: Letter Names (0-52): Spring** | | |
| Mean (SD) | 34.5 (10.3) | 34.0 (10.9) |
| Median [Min, Max] | 36.2<br>[0, 52.0] | 35.4<br>[4.30, 52.0] |
| Missing | 15347 (93.1%) | 1919 (88.1%) |
| **PALS: Letter Sounds (0-25): Spring** | | |
| Mean (SD) | 12.5 (5.98) | 12.6 (6.22) |

| | | |
|---|---|---|
| Median [Min, Max] | 12.7<br>[0, 25.0] | 13.8<br>[0, 25.0] |
| Missing | 15347 (93.1%) | 1919 (88.1%) |
| **PALS: Name writing (0-5): Spring** | | |
| Mean (SD) | 4.04 (0.974) | 3.96 (1.05) |
| Median [Min, Max] | 4.38<br>[0, 5.00] | 4.39<br>[1.00, 5.00] |
| Missing | 15350 (93.1%) | 1919 (88.1%) |
| **PALS: Passage Comprehension: Expressive (0-5): Spring** | | |
| Mean (SD) | 2.54 (0.906) | 2.60 (0.863) |
| Median [Min, Max] | 2.59<br>[0, 5.00] | 2.67<br>[0, 4.88] |
| Missing | 15351 (93.1%) | 1920 (88.2%) |
| **PALS: Passage Comprehension: Retell (0-6): Spring** | | |
| Mean (SD) | 2.75 (1.14) | 2.76 (1.06) |
| Median [Min, Max] | 2.86<br>[0, 5.97] | 2.84<br>[0, 5.67] |
| Missing | 15351 (93.1%) | 1920 (88.2%) |
| **PALS: Passage Comprehension: Receptive (0-4): Spring** | | |
| Mean (SD) | 2.70 (0.773) | 2.68 (0.732) |
| Median [Min, Max] | 2.81<br>[0, 4.00] | 2.77<br>[0, 4.00] |
| Missing | 15352 (93.1%) | 1920 (88.2%) |
| **PALS: Syllable Segmenting (0-10): Spring** | | |
| Mean (SD) | 7.52 (2.24) | 7.76 (1.94) |
| Median [Min, Max] | 8.06<br>[0, 10.0] | 8.25<br>[0.167, 10.0] |
| Missing | 15348 (93.1%) | 1919 (88.1%) |

**Table 16**

*Total Days per Week Participants Reported using Different Elements of STREAMin³ based on whether they received coaching or not the month prior to the End of Year Survey*

| | Met With Coach (N=44) | Did Not Meet with Coach (N=16) | Overall (N=60) |
|---|---|---|---|
| **Provocation** | | | |
| 0 Days | 4 (9.1%) | 6 (37.5%) | 10 (16.7%) |
| 1 Day | 5 (11.4%) | 3 (18.8%) | 8 (13.3%) |
| 2 Days | 2 (4.5%) | 1 (6.3%) | 3 (5.0%) |
| 3 Days | 2 (4.5%) | 2 (12.5%) | 4 (6.7%) |
| 4-5 Days | 31 (70.4%) | 4 (25.0%) | 35 (58.3%) |
| **Core Skills** | | | |
| 0 Days | 0 (0%) | 2 (12.5%) | 2 (3.3%) |
| 1 Day | 6 (13.6%) | 6 (37.5%) | 12 (20.0%) |
| 2 Days | 2 (4.5%) | 2 (12.5%) | 4 (6.7%) |
| 3 Days | 4 (9.1%) | 0 (0%) | 4 (6.7%) |
| 4 - 5 Days | 32 (72.7%) | 6 (37.5%) | 38 (63.3%) |
| **Group Activity** | | | |
| 0 Days | 4 (9.1%) | 6 (37.5%) | 10 (16.7%) |
| 1 Day | 2 (4.5%) | 3 (18.8%) | 5 (8.3%) |
| 2 Days | 8 (18.2%) | 4 (25.0%) | 12 (20.0%) |
| 3 Days | 4 (9.1%) | 1 (6.3%) | 5 (8.3%) |
| 4 - 5 Days | 26 (59.1%) | 2 (12.6%) | 28 (46.7%) |
| **Games** | | | |
| 0 Days | 29 (65.9%) | 10 (62.5%) | 39 (65.0%) |
| 1 Day | 5 (11.4%) | 3 (18.8%) | 8 (13.3%) |
| 2 Days | 2 (4.5%) | 0 (0%) | 2 (3.3%) |

| | | | |
|---|---|---|---|
| 3 Days | 1 (2.3%) | 2 (12.5%) | 3 (5.0%) |
| 4 - 5 Days | 7 (15.9%) | 1 (6.3%) | 8 (13.4%) |
| **Activity Cards** | | | |
| 0 Days | 7 (15.9%) | 7 (43.8%) | 14 (23.3%) |
| 1 Day | 4 (9.1%) | 4 (25.0%) | 8 (13.3%) |
| 2 Days | 4 (9.1%) | 1 (6.3%) | 5 (8.3%) |
| 3 Days | 6 (13.6%) | 2 (12.5%) | 8 (13.3%) |
| 4 - 5 Days | 23 (52.3%) | 2 (12.5%) | 25 (41.7%) |

**Table 17**

*Total Days per Week Participants Reported using Different Elements of STREAMin³ based on how much coaching they received the month prior to the End of Year Survey*

| | Less than 30 minutes (N=7) | 30-60 minutes (N=20) | 1-1.5 hours (N=8) | 1.5-2 hours (N=5) | Over 2 hours (N=4) | Overall (N=44) |
|---|---|---|---|---|---|---|
| **Provocation** | | | | | | |
| 0 Days | 0 (0%) | 2 (10.0%) | 0 (0%) | 2 (40.0%) | 0 (0%) | 4 (9.1%) |
| 1 Day | 1 (14.3%) | 1 (5.0%) | 3 (37.5%) | 0 (0%) | 0 (0%) | 5 (11.4%) |
| 2 Days | 0 (0%) | 1 (5.0%) | 1 (12.5%) | 0 (0%) | 0 (0%) | 2 (4.5%) |
| 3 Days | 0 (0%) | 2 (10.0%) | 0 (0%) | 0 (0%) | 0 (0%) | 2 (4.5%) |
| 4-5 Days | 6 (85.7%) | 14 (70.0%) | 4 (50.0%) | 3 (60.0%) | 4 (100.0%) | 31 (70.4%) |
| **Core Skills** | | | | | | |
| 0 Days | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| 1 Day | 0 (0%) | 3 (15.0%) | 2 (25.0%) | 0 (0%) | 1 (25.0%) | 6 (13.6%) |
| 2 Days | 0 (0%) | 1 (5.0%) | 1 (12.5%) | 0 (0%) | 0 (0%) | 2 (4.5%) |
| 3 Days | 0 (0%) | 3 (15.0%) | 1 (12.5%) | 0 (0%) | 0 (0%) | 4 (9.1%) |
| 4-5 Days | 12 (100.0%) | 13 (65.0%) | 4 (50.0%) | 5 (100.0%) | 3 (75.0%) | 32 (72.7%) |
| **Group Activity** | | | | | | |
| 0 Days | 1 (14.3%) | 1 (5.0%) | 0 (0%) | 1 (20.0%) | 1 (25.0%) | 4 (9.1%) |
| 1 Days | 0 (0%) | 0 (0%) | 0 (0%) | 2 (40.0%) | 0 (0%) | 2 (4.5%) |
| 2 Days | 0 (0%) | 6 (30.0%) | 2 (25.0%) | 0 (0%) | 0 (0%) | 8 (18.2%) |
| 3 Days | 1 (14.3%) | 1 (5.0%) | 1 (12.5%) | 0 (0%) | 1 (25.0%) | 4 (9.1%) |
| 4- 5 Days | 5 (71.5%) | 12 (60.0%) | 5 (62.5%) | 2 (40.0%) | 2 (50.0%) | 26 (59.1%) |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Games** | | | | | | |
| 0 Days | 5 (71.4%) | 13 (65.0%) | 4 (50.0%) | 4 (80.0%) | 3 (75.0%) | 29 (65.9%) |
| 1 Day | 2 (28.6%) | 1 (5.0%) | 2 (25.0%) | 0 (0%) | 0 (0%) | 5 (11.4%) |
| 2 Days | 0 (0%) | 2 (10.0%) | 0 (0%) | 0 (0%) | 0 (0%) | 2 (4.5%) |
| 3 Days | 0 (0%) | 0 (0%) | 1 (12.5%) | 0 (0%) | 0 (0%) | 1 (2.3%) |
| 4- 5 Days | 0 (0%) | 4 (20.0%) | 1 (12.5%) | 1 (20.0%) | 1 (25.0%) | 7 (15.9%) |
| **Activity Cards** | | | | | | |
| 0 Days | 0 (0%) | 4 (20.0%) | 2 (25.0%) | 0 (0%) | 1 (25.0%) | 7 (15.9%) |
| 1 Day | 1 (14.3%) | 2 (10.0%) | 1 (12.5%) | 0 (0%) | 0 (0%) | 4 (9.1%) |
| 2 Days | 1 (14.3%) | 1 (5.0%) | 0 (0%) | 2 (40.0%) | 0 (0%) | 4 (9.1%) |
| 3 Days | 1 (14.3%) | 2 (10.0%) | 2 (25.0%) | 1 (20.0%) | 0 (0%) | 6 (13.6%) |
| 4- 5 Days | 4 (57,2%) | 11 (55.0%) | 3 (37.5%) | 2 (40.0%) | 3 (75.0%) | 23 (52.3%) |

.

**Table 18**

*Regression Analysis Summary for Variables Predicting Dosage of Delivery: Provocation Total (n = 42)*

| Predictor | B | SE | *t* | *p* |
|---|---|---|---|---|
| Intercept | 2.55 | 1.03 | 2.47 | .018* |
| Current Position | | | | |
| Pre-K Teacher | -0.40 | 0.57 | -0.70 | .487 |
| Teaching License | | | | |
| Yes | -1.14 | 0.74 | -1.53 | .134 |
| Education | | | | |
| Bachelor's Degree | -1.24 | 0.82 | -1.50 | .141 |
| Education Specialist | 2.20 | 2.08 | 1.06 | .296 |
| High School Diploma | -0.43 | 0.72 | -0.59 | .556 |
| Master's Degree | -2.98 | 1.49 | -2.00 | .052 |
| Other | 1.19 | 1.38 | 0.87 | .392 |
| Years of Experience | | | | |
| 6-10 years | 0.66 | 0.74 | 0.88 | .382 |
| 11-15 years | -0.46 | 0.99 | -0.47 | .643 |
| 16-20 years | 1.27 | 0.87 | 1.47 | .150 |
| 20+ years | 0.71 | 1.17 | 0.61 | .548 |
| Met Coach After Post | | | | |
| Yes | 1.60 | 0.67 | 2.39 | .022* |
| Model Fit Information | | | | |
| Residual SE | 1.85 | | | |
| $R^2$ | 0.363 | | | |
| Adjusted $R^2$ | 0.180 | | | |
| *F* | 1.99 | | | |
| df | 12, 42 | | | |
| *p* | 0.05 | | | |

*Note. * p < .05.*

**Table 19**

*Regression Analysis Summary for Variables Predicting Dosage of Delivery: Core Skills Total (n = 42)*

| Predictor | B | SE | *t* | *p* |
|---|---|---|---|---|
| Intercept | 3.00 | 0.82 | 3.68 | <.001*** |
| Current Position | | | | |
| Pre-K Teacher | -0.38 | 0.45 | -0.84 | .406 |
| Teaching License | | | | |
| Yes | -0.79 | 0.59 | -1.34 | .187 |

| | | | | |
|---|---|---|---|---|
| Education | | | | |
| Bachelor's Degree | -2.01 | 0.65 | -3.08 | .004** |
| Education Specialist | 0.62 | 1.64 | 0.38 | .710 |
| High School Diploma | -0.26 | 0.57 | -0.46 | .647 |
| Master's Degree | -1.98 | 1.18 | -1.68 | .101 |
| Other | 1.47 | 1.09 | 1.35 | .184 |
| Years of Experience | | | | |
| 6-10 years | 0.76 | 0.59 | 1.30 | .202 |
| 11-15 years | 0.60 | 0.78 | 0.77 | .448 |
| 16-20 years | 0.01 | 0.69 | 0.01 | .991 |
| 20+ years | -0.39 | 0.93 | -0.42 | .678 |
| Met Coach After Post | | | | |
| Yes | 1.57 | 0.53 | 2.97 | .005** |
| Model Fit Information | | | | |
| Residual SE | 1.47 | | | |
| $R^2$ | 0.450 | | | |
| Adjusted $R^2$ | 0.293 | | | |
| $F$ | 2.86 | | | |
| df | 12, 42 | | | |
| $p$ | 0.0058 | | | |

Notes. ** $p < .01$. *** $p < .001$.

**Table 20**

*Regression Analysis Summary for Variables Predicting Dosage of Delivery: Group Activity (n = 42)*

| Predictor | B | SE | *t* | *p* |
|---|---|---|---|---|
| Intercept | 0.86 | 0.92 | 0.94 | .355 |
| Current Position | | | | |
| Pre-K Teacher | 0.57 | 0.51 | 1.11 | .272 |
| Teaching License | | | | |
| Yes | -0.65 | 0.66 | -0.99 | .329 |
| Education | | | | |
| Bachelor's Degree | -0.13 | 0.73 | -0.18 | .855 |
| Education Specialist | -1.11 | 1.84 | -0.60 | .550 |
| High School Diploma | -0.21 | 0.64 | -0.33 | .741 |
| Master's Degree | 1.91 | 1.32 | 1.44 | .157 |
| Other | 2.56 | 1.22 | 2.09 | .042* |
| Years of Experience | | | | |
| 6-10 years | 0.69 | 0.66 | 1.04 | .303 |
| 11-15 years | 0.06 | 0.88 | 0.07 | .943 |
| 16-20 years | 0.53 | 0.77 | 0.69 | .497 |
| 20+ years | -0.73 | 1.04 | -0.70 | .490 |
| Met Coach After Post | | | | |
| Yes | 2.22 | 0.59 | 3.73 | <.001*** |
| Model Fit Information | | | | |
| Residual SE | 1.64 | | | |
| $R^2$ | 0.399 | | | |
| Adjusted $R^2$ | 0.227 | | | |
| *F* | 2.33 | | | |
| df | 12, 42 | | | |
| *p* | 0.022 | | | |

Notes. * $p < .05$. *** $p < .001$.

.

**Table 21**

*Regression Analysis Summary for Variables Predicting Dosage of Delivery: Games (n = 42)*

| Predictor | B | SE | *t* | *p* |
|---|---|---|---|---|
| Intercept | -0.36 | 0.88 | -0.41 | .681 |
| Current Position | | | | |
| Pre-K Teacher | 1.62 | 0.49 | 3.31 | .0019** |
| Teaching License | | | | |
| Yes | -0.02 | 0.63 | -0.04 | .971 |
| Education | | | | |
| Bachelor's Degree | -0.08 | 0.70 | -0.11 | .910 |
| Education Specialist | 2.71 | 1.77 | 1.54 | .132 |
| High School Diploma | 0.55 | 0.61 | 0.89 | .377 |
| Master's Degree | -1.05 | 1.27 | -0.83 | .412 |
| Other | -1.10 | 1.17 | -0.94 | .355 |
| Years of Experience | | | | |
| 6-10 years | 1.03 | 0.63 | 1.63 | .112 |
| 11-15 years | -0.24 | 0.84 | -0.28 | .778 |
| 16-20 years | 0.13 | 0.74 | 0.18 | .861 |
| 20+ years | 1.43 | 1.00 | 1.43 | .161 |
| Met Coach After Post | | | | |
| Yes | 0.18 | 0.57 | 0.32 | .754 |
| Model Fit Information | | | | |
| Residual SE | 1.58 | | | |
| $R^2$ | 0.381 | | | |
| Adjusted $R^2$ | 0.204 | | | |
| *F* | 2.16 | | | |
| df | 12, 42 | | | |
| *p* | 0.033 | | | |

Notes. ** $p < .01$.

**Table 22**

*Regression Analysis Summary for Variables Predicting Dosage of Delivery: Activity Cards (n = 42)*

| Predictor | B | SE | *t* | *p* |
|---|---|---|---|---|
| Intercept | 1.88 | 0.95 | 1.98 | .054 |
| Current Position | | | | |
| Pre-K Teacher | -1.27 | 0.53 | -2.41 | .020* |
| Teaching License | | | | |
| Yes | -0.52 | 0.68 | -0.76 | .454 |
| Education | | | | |
| Bachelor's Degree | -0.78 | 0.76 | -1.02 | .312 |
| Education Specialist | 3.45 | 1.91 | 1.81 | .078 |
| High School Diploma | 0.17 | 0.66 | 0.26 | .797 |
| Master's Degree | -2.63 | 1.37 | -1.92 | .061 |
| Other | -0.12 | 1.26 | -0.10 | .923 |
| Years of Experience | | | | |
| 6-10 years | 0.95 | 0.68 | 1.39 | .172 |
| 11-15 years | -0.81 | 0.91 | -0.90 | .375 |
| 16-20 years | -0.00 | 0.80 | -0.01 | .996 |
| 20+ years | 0.71 | 1.08 | 0.66 | .512 |
| Met Coach After Post | | | | |
| Yes | 1.83 | 0.61 | 2.98 | .005** |
| Model Fit Information | | | | |
| Residual SE | 1.62 | | | |
| $R^2$ | 0.378 | | | |
| Adjusted $R^2$ | 0.201 | | | |
| *F* | 2.13 | | | |
| df | 12, 42 | | | |
| *p* | 0.035 | | | |

Notes. * $p < .05$. ** $p < .01$.