HARVARD

GRADUATE SCHOOL OF EDUCATION

Experimental Evidence on the Impact of Tutoring Format and Tutors: Findings from an Early Literacy Program

Shirin Hashim, Harvard University

NSSA Conference

Research Lightening Round

May 9, 2024

CUNY Reading Corps

- **Tutors**: undergraduate and graduate CUNY preservice teachers
- **Students**: 1^{st}-3^{rd} grade striving readers
- **Programs**: Reading Go! Reading Ready
- **Implementation**:
  - Primarily 1-to-1
  - 3-5 sessions/week
  - 30-45 minutes/session
  - In-person and remote
  - Formative and summative assessments
  - Tutor training + on-going supports
  - Consistent tutor throughout

Virtual tutoring shows potential

Online Tutoring by College Volunteers: Experimental Evidence from a Pilot Program

Matthew A. Kraft
John A. List
Jeffrey A. Livingston
Sally Sadoff

The Effects of Virtual Tutoring on Young Readers: Results from a Randomized Controlled Trial

Carly D. Robinson, Stanford University
Cynthia Pollard, UnboundEd
Sarah Novicoff, Stanford University
Sara White, Vanderbilt University
Susanna Loeb, Stanford University

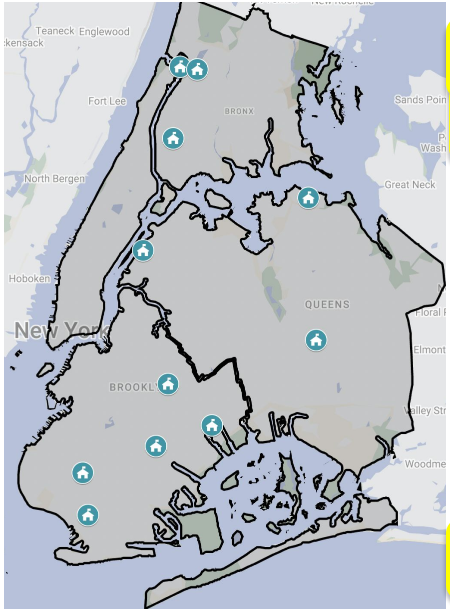
The Effects of In-School Virtual Tutoring on Student Reading Development: Evidence from a Short-Cycle Randomized Controlled Trial

Douglas D. Ready, Sierra G. McCormick, Rebecca J. Shmoys

Summer 2023: 6 weeks of daily tutoring

CUNY Reading Corps

# Tutors	139
Female	79%
Asian	36%
Black	16%
Hispanic	27%
White	14%
Ed-Related Major	47%
Avg. Cum. GPA	3.39



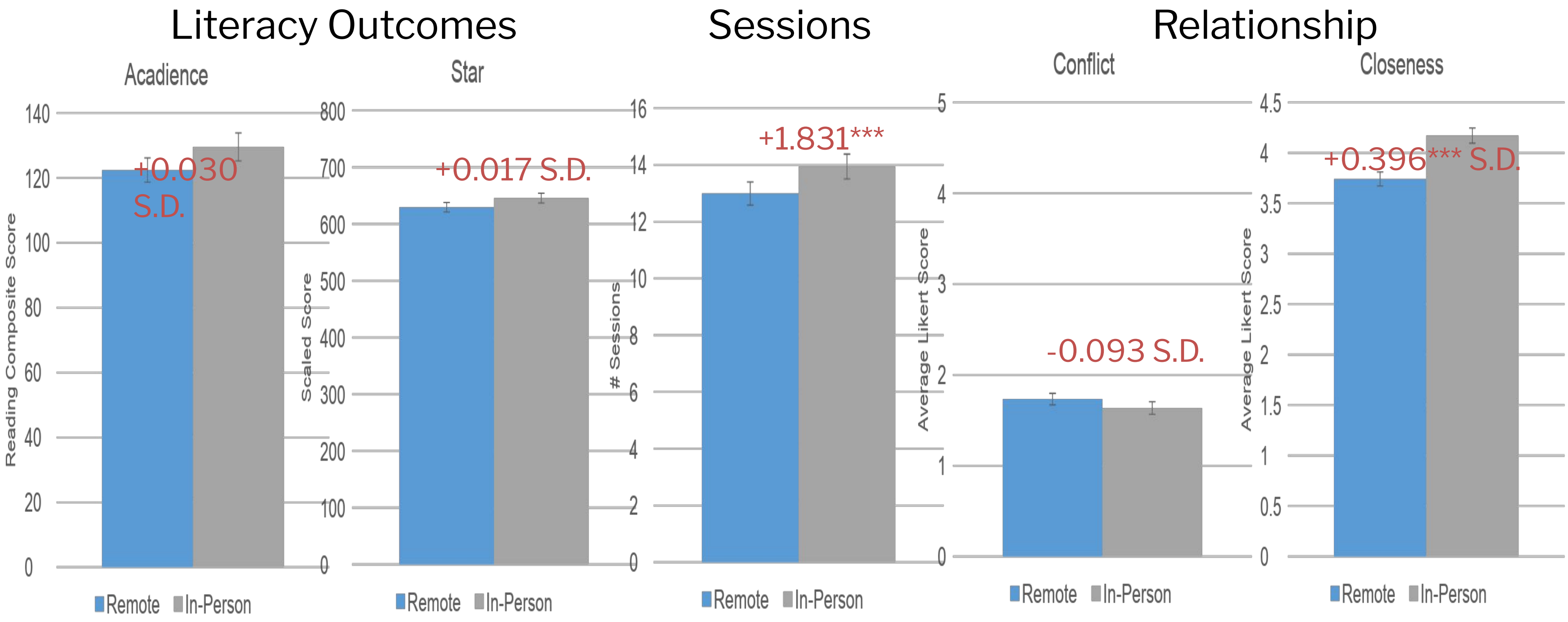
NYC Public Schools

# 1st-3rd Grade Students	607
In-Person (vs. Remote)	45%
Asian	23%
Black	28%
Hispanic	36%
White	10%
English Language Learner	37%
Students with Disabilities	24%
Economically Disadvantaged	83%

Student outcomes

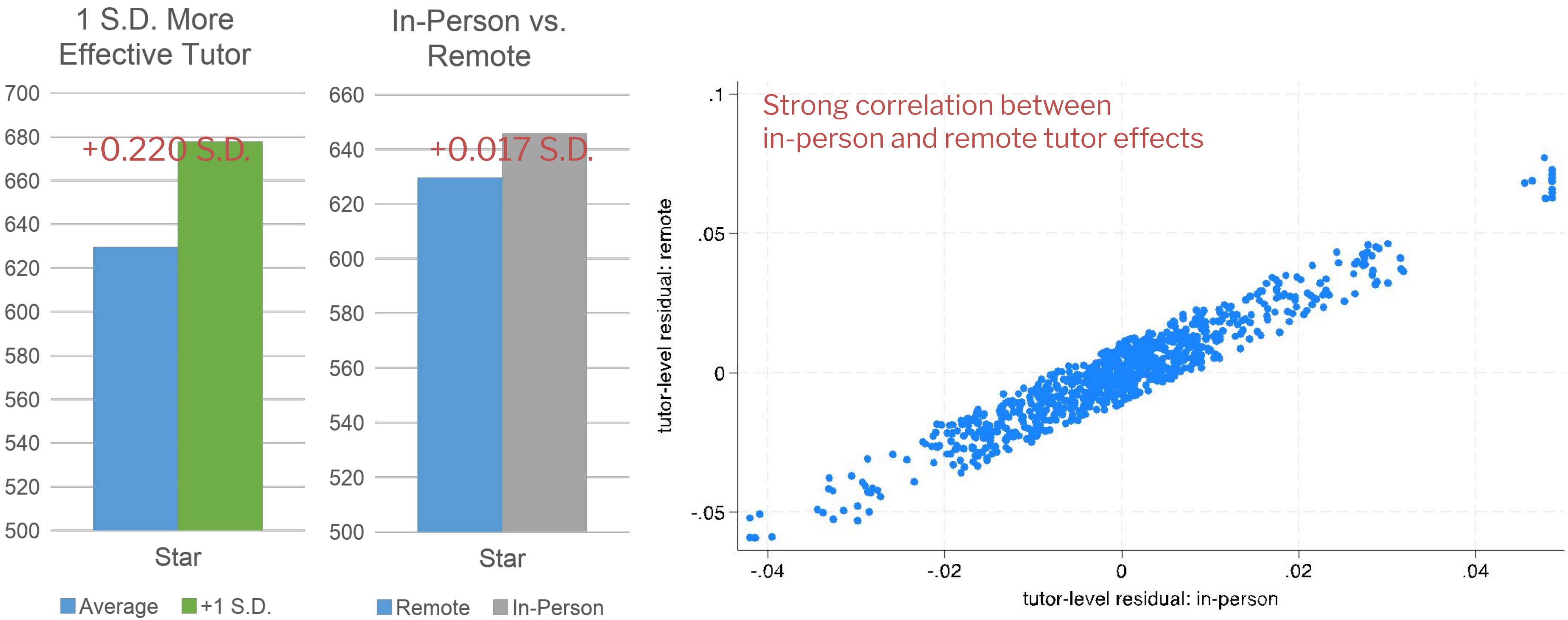
- **Literacy Outcomes:**
    - Star Early Literacy standardized scaled scores
    - Acadience Reading standardized composite scores
- **Student-Tutor Relationship:** (Pianta, 1992)
    - Closeness (feelings of affection and open communication)
    - Conflict (feelings of negativity and conflict)
- **Attendance:**
    - Number of tutoring sessions

Impact instructional format depends on outcome



Bar graphs display average outcomes across all sites for students with average pre-test scores and with an average tutor. Differences between in-person and remote additionally control for student demographics, differential attrition, and differences between sites by grade.

Large differences between tutors, but not by format



Bar graphs display average outcomes across all sites for students with average pre-test scores. The tutor effects bar graph holds constant the format of instruction (remote). The instructional format bar graph holds constant tutor efficacy (average tutor.)Differences in standard deviation units additionally control for student demographics, differential attrition, and differences between sites by grade.

Some tutor characteristics impact literacy outcomes

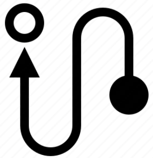
Impact estimates (in S.D. units)

	Acadience	Star	Number of Sessions	Conflict	Closeness
GPA	-0.046 (0.049)	0.103 (0.080)	0.758 (0.612)	-0.074 (0.098)	0.133 (0.113)
Education-Related Major	0.128** (0.059)	0.156* (0.092)	0.952 (0.753)	-0.084 (0.123)	0.079 (0.145)
Bachelor's Degree (vs. Associate's)	0.059 (0.076)	-0.072 (0.134)	-0.198 (0.982)	0.000 (0.158)	-0.200 (0.188)

"Science of Reading" effect?

Take-aways

1. Remote tutoring could be a feasible option
   - But details matter:
     - Materials designed for both modalities
     - Site coordinators
     - In-school tutoring
     - Tutoring in both formats
     - High-impact



2. Consider indirect impacts (attendance & relationships) and potential downstream effects on literacy

3. Identify, support, retain effective tutors
   - Formal training, other characteristics, observations

THANK YOU

Katie Pace Miles, Ph.D.

Assoc. Professor of Early Literacy

Brooklyn College, CUNY

Co-Founder & P.I.

CUNY Reading Corps

katiepacemilesphd.com

Shirin Hashim, Ph.D.

Harvard Graduate School of Education

shirin_hashim@g.harvard.edu

Erin Croke, Ph.D.

Director of Literacy Initiatives

CUNY Reading Corps

CUNY Central Office of Academic Affairs

cunyreadingcorps.edu

CUNY Career Launch · Fund for Public Schools · Anonymous Family Foundation · Altman Foundation · Benedict Silverman Foundation · Heckscher Foundation for Children · Robin Hood · Tiger Foundation · NYC Young Men's Initiative · Summer Youth Employment Program

Appendix

What do effective tutoring programs look like?

Skilled in relationship-building, content, cultural competence
Consistent
Trained, receive on-going coaching, held accountable
Reflective of student population

Supplemental, high-quality, aligned materials
3+ sessions/week
30+ minutes/session
1-3 students/tutor

Embedded in school
Family engagement

Formative and on-going assessment to tailor instruction
Summative assessment to assess program efficacy

Adapted from "Tutoring Quality Standards," NSSA (2023)

Research questions

Students randomly assigned to format

RQ1: What is the impact of instructional modality (in-person versus remote) on student outcomes?

Tutors randomly assigned to both formats

RQ2: How much of the overall variation in student outcomes can be attributed to differences between tutors?

Access to multiple student outcomes and administrative data

RQ3: How do effects vary based on tutor characteristics, instructional modality, and across different outcomes?

Balance Table

	(1)	(2)	(3)	(4)	(5)	(6)
	Acadience Sample (N=446)			Star Sample (N=426)		
	In-Person	Remote	*p-value*	In-Person	Remote	*p-value*
N	186	260	--	210	216	--
Rising 1st Grade	0.452	0.538	--	0.352	0.472	--
Rising 2nd Grade	0.548	0.462	--	0.419	0.431	--
Rising 3rd Grade	--	--	--	0.229	0.097	--
Avg. Standardized Pre-Test	-0.196	-0.207	0.975	0.024	-0.062	0.525
Missing STRS Response	0.151	0.196	0.171	0.167	0.102	0.187
Has Admin. Data	1.000	1.000	--	0.962	0.954	0.080*
Amongst those with data:						
Female	0.511	0.438	0.014**	0.520	0.417	0.008***
Asian	0.242	0.269	0.215	0.238	0.277	0.193
Black	0.290	0.285	0.582	0.262	0.257	0.496
Hispanic	0.333	0.296	0.759	0.406	0.311	0.875
White	0.091	0.131	0.471	0.074	0.150	0.313
ELL	0.328	0.431	0.089*	0.332	0.417	0.225
ED	0.882	0.812	0.270	0.856	0.816	0.980
SWD	0.220	0.215	0.920	0.252	0.223	0.789

Attrition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Randomized Sample (N=939)			Participant Sample (N=607)			Missing Acadience (N=73)			Missing Star (N=181)		
	In-Person	Remote	*p-value*	In-Person	Remote	*p-value*	In-Person	Remote	*p-value*	In-Person	Remote	*p-value*
N	401	538	--	273	334	--	32	41	--	63	118	--
Withdrawal (from program)	0.319	0.379	0.034**	--	--	--	--	--	--	--	--	--
Rising 1st Grade	0.359	0.448	--	0.374	0.488	--	0.562	0.561	--	0.444	0.517	--
Rising 2nd Grade	0.416	0.390	--	0.425	0.413	--	0.438	0.439	--	0.444	0.381	--
Rising 3rd Grade	0.224	0.162	--	0.201	0.099	--	--	--	--	0.111	0.102	--
Missing Star Pre-Test	--	--	--	0.011	0.054	0.005***	--	--	--	0.048	0.153	0.003***
Missing Star Post-Test	--	--	--	0.231	0.353	0.003***	--	--	--	1.000	1.000	--
Avg. Std. Star Pre-Test	--	--	--	0.020	-0.070	0.461	--	--	--	0.007	-0.086	0.552
Missing Spring Acadience	--	--	--	0.092	0.073	0.943	0.625	0.537	0.336	--	--	--
Missing Fall Acadience	--	--	--	0.128	0.123	0.448	0.875	0.902	0.186	--	--	--
Avg. Std. Spring Acadience	--	--	--	-0.281	-0.255	0.972	-0.804	-0.325	0.455	--	--	--
Missing STRS Response	--	--	--	0.183	0.195	0.397	0.312	0.171	0.829	0.238	0.364	0.020**
Has Admin. Data	0.958	0.950	0.044**	0.960	0.958	0.067*	0.781	0.756	0.789	0.952	0.966	0.812
Amongst those with data:												
Female	0.487	0.456	0.160	0.504	0.438	0.006***	0.480	0.355	0.903	0.450	0.474	0.681
Asian	0.180	0.190	0.137	0.214	0.250	0.330	0.080	0.065	0.494	0.133	0.202	0.600
Black	0.271	0.335	0.385	0.271	0.291	0.458	0.320	0.355	0.954	0.300	0.351	0.933
Hispanic	0.430	0.339	0.851	0.401	0.322	0.864	0.560	0.484	0.954	0.383	0.342	0.448
White	0.081	0.114	0.244	0.084	0.122	0.578	0.040	0.097	0.494	0.117	0.070	0.519
ELL	0.312	0.356	0.068*	0.324	0.403	0.138	0.400	0.226	0.747	0.300	0.377	0.226
ED	0.857	0.826	0.722	0.866	0.806	0.395	0.760	0.774	0.491	0.900	0.789	0.238
SWD	0.258	0.243	0.851	0.256	0.228	0.839	0.120	0.226	0.106	0.267	0.237	0.678

Descriptive Statistics

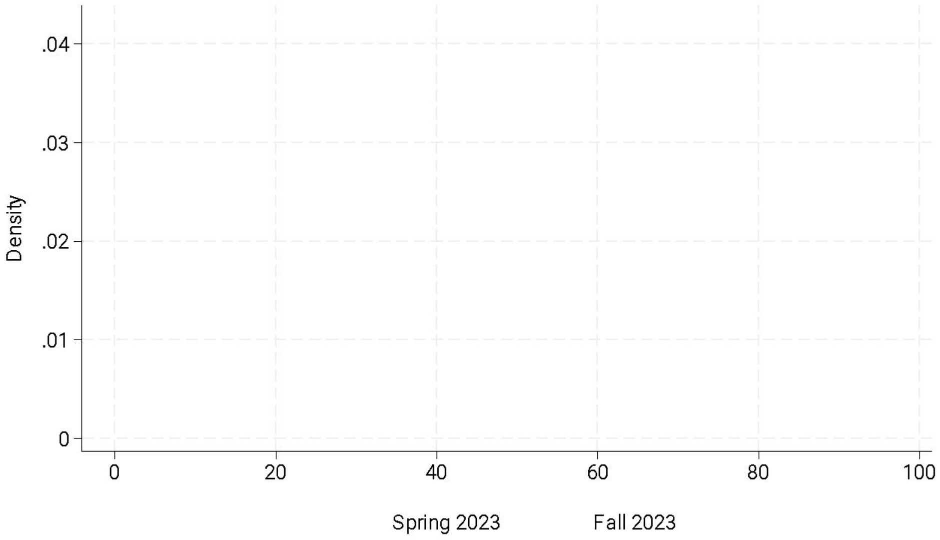
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Total			In-Person			Remote		
	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
Acadience Reading:									
Spring 2023 RCS	113.657	84.475	446	116.022	87.551	186	111.965	82.333	260
Fall 2023 RCS	112.303	75.372	446	119.048	79.483	186	107.477	72.056	260
Spring 2023 % Well Below	0.439	--	446	0.435	--	186	0.442	--	260
Fall 2023 % Well Below	0.469	--	446	0.430	--	186	0.481	--	260
Star Early Literacy:									
Pre-Test SS	630.211	133.131	426	643.086	132.652	210	617.694	132.706	216
Post-Test SS	634.190	133.130	426	648.871	140.204	210	619.917	136.023	216
Pre-Test NCE	28.248	21.609	426	28.016	20.439	210	28.473	22.735	216
Post-Test NCE	31.073	21.809	426	31.204	21.166	210	30.946	22.466	216
Non-Academic:									
Number of Sessions	13.114	6.424	607	13.949	6.331	273	12.431	6.332	334
Conflict Score	1.709	0.989	492	1.636	1.013	223	1.769	0.967	269
Closeness Score	3.911	0.957	492	4.133	0.847	223	3.727	1.004	269

Pairwise Correlations between Outcomes

	(1)	(2)	(3)	(4)	(5)
	Acadience	Star	Number of Sessions	Conflict	Closeness
Acadience	1.000	0.543	0.051	-0.143	0.103
Star	0.543	1.000	0.110	-0.265	0.120
Number of Sessions	0.051	0.110	1.000	0.067	0.048
Conflict	-0.143	-0.265	0.067	1.000	-0.391
Closeness	0.103	0.120	0.048	-0.391	1.000

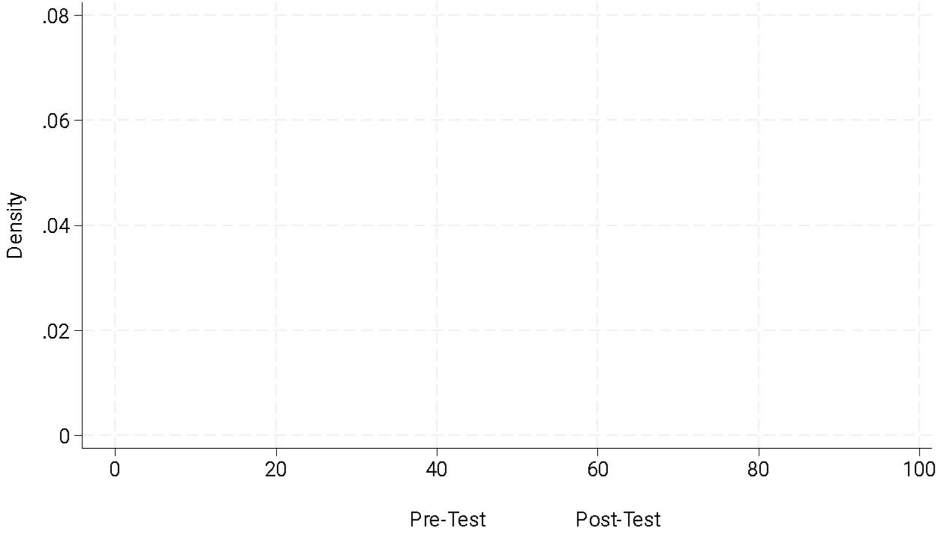
Pre-Post Distribution of Percentile Ranks

Acadience Reading



Notes: Data are shown for students in the analytic sample who had complete pre-post data for the Acadience Reading assessment (N=446).

Star Early Literacy



Notes: Data are shown for students in the analytic sample who had complete pre-post data for the Star Early Literacy assessment (N=426).

Impact of In-Person over Remote Tutoring

RQ1: What is the impact of instructional modality (in-person versus remote) on students' literacy outcomes, attendance, and tutor-student relationships?

	(1) Acadience	(2) Star	(3) Sessions	(4) Conflict	(5) Closeness
Fixed parameters:					
β_1 (in-person)	0.030	0.017	1.831***	-0.093	0.396***
	(0.050)	(0.082)	(0.532)	(0.099)	(0.092)
β_0 (constant)	0.308***	0.352*	11.872***	0.639**	-0.172
	(0.130)	(0.202)	(1.460)	(0.276)	(0.242)
Random parameters:					
σ_{u0} (tutor intercept s.d.)	0.060	0.220	2.234	0.191	0.496
	(0.068)	(0.064)	(0.352)	(0.101)	(0.058)
σ_e (student-level residual s.d.)	0.467	0.683	5.131	0.912	0.718
	(0.018)	(0.029)	(0.177)	(0.035)	(0.028)
Number of Students	446	408	565	467	467
Number of Tutors	129	127	137	117	117

Tutor-Level Random Effects

RQ2: How much of the overall variation in student outcomes can be attributed to differences between tutors?

	(1) Acadience	(2) Star	(3) Sessions	(4) Conflict	(5) Closeness
Fixed parameters:					
β_1 (in-person)	0.030	0.017	1.831***	-0.093	0.396***
	(0.050)	(0.082)	(0.532)	(0.099)	(0.092)
β_0 (constant)	0.308***	0.352*	11.872***	0.639**	-0.172
	(0.130)	(0.202)	(1.460)	(0.276)	(0.242)
Random parameters:					
σ_{u0} (tutor intercept s.d.)	0.060	0.220	2.234	0.191	0.496
	(0.068)	(0.064)	(0.352)	(0.101)	(0.058)
σ_e (student-level residual s.d.)	0.467	0.683	5.131	0.912	0.718
	(0.018)	(0.029)	(0.177)	(0.035)	(0.028)
Number of Students	446	408	565	467	467
Number of Tutors	129	127	137	117	117

Heterogeneity by Tutor Characteristics

RQ3: How do effects vary based on *tutor characteristics*?

	(1)	(2)	(3)	(4)	(5)
	Acadience	Star	Sessions	Conflict	Closeness
Fixed parameters:					
β_1 (in-person)	0.036	0.025	1.785***	-0.091	0.413***
	(0.053)	(0.083)	(0.581)	(0.110)	(0.100)
β_2 (cumulative GPA)	-0.046	0.103	0.758	-0.074	0.133
	(0.049)	(0.080)	(0.612)	(0.098)	(0.113)
β_3 (education major)	0.128**	0.156*	0.952	-0.084	0.079
	(0.059)	(0.092)	(0.753)	(0.123)	(0.145)
β_4 (bachelor's degree)	0.059	-0.072	-0.198	0.000	-0.200
	(0.076)	(0.134)	(0.982)	(0.158)	(0.188)
β_0 (constant)	0.319	0.617	8.768***	1.013**	-0.370
	(0.220)	(0.429)	(2.618)	(0.461)	(0.472)
Random parameters:					
σ_{u0} (tutor intercept s.d.)	0.067	0.148	2.294	0.108	0.477
	(0.066)	(0.086)	(0.373)	(0.186)	(0.061)
σ_e (student-level residual s.d.)	0.470	0.653	5.133	0.933	0.712
	(0.019)	(0.030)	(0.190)	(0.039)	(0.030)

Tutor Random Effects for In-Person and Remote

RQ3: How do effects vary based on *instructional modality*?

	(1)	(2)	(3)	(4)	(5)
	Acadience	Star	Sessions	Conflict	Closeness
σ_{u0} (tutor remote intercept s.d.)	0.071	0.222	2.431	0.363	0.509
σ_{u1} (tutor in-person intercept s.d.)	0.050	0.217	2.632	0.054	0.478
σ_e (student-level residual s.d.)	0.466	0.683	5.005	0.892	0.718
$\rho_{u1,u0}$ (corr. remote, in-person)	1.000	1.000	0.540	1.000	1.000
Number of Students	446	408	565	467	467
Number of Tutors	129	127	137	117	117

Correlations of Tutor Effects Between Outcomes

RQ3: How do effects vary across *different outcomes*?

	(1)	(2)	(3)	(4)	(5)
	Acadience	Star	Number of Sessions	Conflict	Closeness
Acadience	1.000	1.000	--	--	0.092
Star	1.000	1.000	0.014	-0.100	0.005
Number of Sessions	--	0.014	1.000	0.306	0.169
Conflict	--	-0.100	0.306	1.000	-0.188
Closeness	0.092	0.005	0.169	-0.188	1.000

Relationship Survey

	Original Version: Student-Teacher Relationship Scale - Short Form (Pianta, 1992)	Modified Version: Student-Tutor Relationship Scale
Response Categories	1 - Definitely does not apply 2 - Not really 3 - Neutral, not sure 4 - Applies somewhat 5 - Definitely applies	1 - Definitely does not apply 2 - Not really 3 - Neutral, not sure 4 - Applies somewhat 5 - Definitely applies
Questions	1. I share an affectionate, warm relationship with this child. 2. This child and I always seem to be struggling with each other. 3. If upset, this child will seek comfort from me. 4. This child is uncomfortable with physical affection or touch from me. 5. This child values his/her relationship with me. 6. When I praise this child, he/she beams with pride. 7. This child spontaneously shares information about himself/herself. 8. This child easily becomes angry with me. 9. It is easy to be in tune with what this child is feeling. 10. This child remains angry or is resistant after being disciplined. 11. Dealing with this child drains my energy. 12. When this child is in a bad mood, I know we're in for a long and difficult day. 13. This child's feelings toward me can be unpredictable or can change suddenly. 14. This child is sneaky or manipulative with me. 15. This child openly shares his/her feelings and experiences with me.	1. I share an affectionate, warm relationship with this child. 2. This child and I always seem to be struggling with each other. 3. If upset, this child will seek comfort from me. 4. This child values his/her relationship with me. 5. When I praise this child, they beam with pride. 6. This child spontaneously shares information about themselves. 7. This child easily becomes angry with me. 8. It is easy to be in tune with what this child is feeling. 9. Dealing with this child drains my energy. 10. When this child is in a bad mood, I know we're in for a long and difficult session. 11. This child's feelings toward me can be unpredictable or can change suddenly. 12. This child openly shares their feelings and experiences with me.
Scoring	Closeness: 1, 3, 4R, 5, 6, 7, 9, 15 Conflict: 2, 8, 10, 11, 12, 13, 14 Subscale scores are the mean of included items. Item 4 is reverse scored.	Closeness: 1, 3, 4, 5, 6, 8, 12 Conflict: 2, 7, 9, 10, 11 Subscale scores are the mean of included items.

Summary of findings

1. It’s hard to improve literacy in a short time
2. We can rule out big differences between the impact of instructional modality on literacy outcomes
3. But there may still be indirect impacts (attendance & relationships) that impact literacy outcomes downstream
4. The between-tutor variation is large
5. Tutors’ GPAs and degree type do not impact outcomes
6. But being enrolled in an education-related major might matter for literacy outcomes
7. Tutor effects do not vary by instructional modality
8. Tutor effects are not correlated across outcomes