Abstract Title Page
Not included in page count.

Title: Developing Content-Area Academic Literacy: A Randomized Control Trial of the Reading Apprenticeship Improving Secondary Education (RAISE) project

Authors and Affiliations:
Cheri Fancsali, Ph.D. Impaq International
Yasuyo Abe, Ph.D., Impaq International
Mikhail Pyatigorsky, Ph.D., Impaq International

Abstract Body

Background / Context:
Nationally, two-thirds of high school students are unable to read and comprehend complex academic materials, think critically about texts, and synthesize information from multiple sources, or communicate what they have learned (NAEP, 2013). Without a substantial change in their academic literacy, U.S. high school students face continued academic problems in high school and college because they are unable to handle the quantity and complexity of assigned reading (ACT, 2012). The Reading Apprenticeship instructional framework was developed two decades ago to help teachers provide the literacy support students need to be successful readers in the content areas. It has since reached over 100,000 teachers in schools across the country, at the middle school, high school and college levels. In 2010, the program developers received a "Validation" grant from the Department of Education's Investing in Innovation Fund (i3) competition to scale-up and conduct a randomized controlled trial of the intervention through a project called Reading Apprenticeship Improving Secondary Success (RAISE).

RAISE worked with high school teachers in three subject areas: English, history and biology. Previous RCTs have tested the efficacy of the Reading Apprenticeship framework and the professional development model in smaller efforts with more closely monitored implementation. These studies demonstrated strong positive effects on teacher practice resulting from the professional development—most notably, teachers' increased use of reading comprehension strategy instruction, metacognitive inquiry routines, and collaborative learning structures in their classrooms. They also showed positive effects on students' literacy and content-area achievement, motivation, and engagement (Greenleaf et al., 2011a, 2011b; Kemple, 2008; Somers, 2010).

This paper addresses an important problem of policy and practice related to building students' academic literacy, aligning with the conference theme: *Lost in Translation: Building Pathways from Knowledge to Action*. The study described here provided information that was useful and responsive to the developer's needs to understand the efficacy of the intervention at scale, and provides useful knowledge for the field about effective instructional frameworks and teacher professional development.

Research Questions:
This study explored implementation questions as well as mediating impact and longer-term impact on student achievement. Questions included:

- To what extent is RAISE implemented in a way that is consistent with the program model and underlying theory of action?
- What are the effects of RAISE on teacher practices and teacher attitudes?
- What are the effects of implementing RAISE on student engagement, and reading attitudes and behaviors?
- What are the effects of implementing RAISE on student literacy achievement?

Setting:
The RCT took place in 42 high schools in California and Pennsylvania. Professional development was provided to teachers in a central location within each state. Teachers were offered on-site support through monthly meetings led by teacher leaders.

Participants:
Prior to random assignment, the study team recruited volunteer teachers from each subject area of interest in each school. To be eligible for the study, teachers needed to teach at least one class in one of the following subject areas: ninth grade ELA, biology, or U.S. history. The random assignment of schools resulted in 22 schools and 130 teachers in the treatment group; 20 schools and 122 teachers in the control group. The student sample consisted of students enrolled in target subject area classes instructed by the participating teachers. In total, three cohorts of students were included in the study, representing teachers' first, second, and third years of Reading Apprenticeship implementation. The student sample included 14,383 students in the first year of implementation, 14,747 in the second year, and 9,194 in the third year.

Exhibits 1 and 2 show the characteristics of teachers and students who participated in the study. There were no statistically significant baseline differences between the treatment and control groups. [Please insert Exhibits 1 & 2 here]

Intervention:
Reading Apprenticeship helps teachers support discipline-specific literacy and learning in their varied content areas by attending to four interacting dimensions of classroom learning culture: Social, Personal, Cognitive, and Knowledge-Building. The **social dimension** involves building community. The classroom becomes a safe environment where students see other students and their teacher as resources for learning. The **personal dimension** includes drawing on students' understandings and experiences as well as developing students' identities as competent readers, building their awareness of their purposes and goals for reading, and connecting current academic tasks to future career or educational goals. The **cognitive dimension** involves developing students' mental processes, including their text-based problem-solving strategies. The **knowledge-building dimension** includes building students' knowledge not only of the content of the text but also of language and word construction, genre and text structure, and discipline-specific discourse practices. At the center of Reading Apprenticeship is an ongoing metacognitive conversation carried on both internally through metacognitive reading and reasoning routines and externally, as teacher and students talk about their personal relationships to reading, the social environment and resources of the classroom, their affective responses and cognitive activity, and the knowledge required to make sense of complex texts. This takes place through extensive reading including increased in-class opportunities for students to practice reading complex academic texts in more skillful ways as they collaborate to make meaning of these texts for learning purposes. The framework targets learning dispositions as well as literacy skills and knowledge. RAISE's intensive professional development included 65 hours of face-to-face sessions delivered over 12 months, with additional in-person, on-site implementation support from teacher leaders.

Research Design:
The study employed a cluster randomized controlled trial to evaluate the efficacy of RAISE. Forty-two schools were randomly assigned to treatment and control conditions. The impact was estimated as "intent-to-treat" effects of the intervention. We estimated a two-level model, with individual students or teachers nested within schools.

Data Collection and Analysis:
The primary sources of data collected and presented in this paper are student record data collected from the district, monthly teacher surveys collected over three years (27 total), student surveys collected at the end of each implementation year and an on-line student literacy assessment developed and collected by ETS at the end of each implementation year (see O'Reilly et. al. 2014 for a description of the assessment and reliability information).

Findings / Results:
Key Findings on Teacher Mediating Outcomes
Monthly teacher surveys measured the extent to which RAISE had an impact on teacher mediating outcomes including changes in instructional practice and confidence in literacy instruction. RAISE had statistically significant impacts on teachers' use of core Reading Apprenticeship practices and on their confidence in delivering literacy instruction with effect sizes ranging from 0.41 to 0.62. The following were areas of impact.

- Fostering student independence
- Students practicing metacognitive conversations
- Students practicing comprehension strategies
- Student collaboration
- Teacher confidence in literacy instruction

- Effects on reported teacher practices were largely driven by large impacts on science teachers.
- The size of the effects increased between year 1 and 2, suggesting that the additional professional development received by teachers in the summer following their first year of implementation, along with the subsequent on-site support during year 2, increased teachers' ability to implement RAISE. See exhibit x. [please insert Exhibit 3 here]

Key Findings on Student Mediating Outcomes
Changes in teacher practices as a result of RAISE are hypothesized to change students' classroom experiences, attitudes and behaviors. These mediating student outcomes were measured through a year-end student survey.

- RAISE produced positive and statistically significant impacts on the full sample of students in two student mediating outcome domains that are hallmarks of the Reading Apprenticeship framework:
    - Increased integration of reading instruction *into* content-area teaching
    - Increased metacognitive inquiry

The effect sizes of the impacts were 0.21 and 0.18 respectively. Impacts in other areas were positive but not statistically significant including outcomes related to collaboration in a community of readers and writers; reader identity; and participation in metacognitive conversations. The size of the impacts on student mediating outcomes increased over time.

There was also a statistically significant impact on participation and contribution to class discussions, and on class time spent reading among science students, and on variety of reading material among history students. The effects on ELA students was smaller and not statistically significant. [Please insert Exhibit 4 here]

Key Findings on Student Achievement Outcomes

Student literacy achievement was measured through an online, scenario-based assessment developed by ETS for this study. The assessment was designed to measure the strategic reading processes that are primary targets of Reading Apprenticeship and closely aligned with the common core state standards. It was designed to be a more rigorous measure of complex reading comprehension than typical state ELA tests.

- By the end of the second year of implementation, RAISE had a positive and statistically significant impact on student literacy in science classes. The effect size of the impact was 0.32.
- Results for the other two subjects were not statistically significant but with a meaningful effect for ELA classrooms (effect size =0.22) and a non-significant result for history classrooms.

The impact in science is particularly interesting given that implementing the Reading Apprenticeship framework may require a more dramatic change in science teachers' core practices and routines than is needed by ELA and history teachers.

- For the full sample and for students in Pennsylvania schools, we found positive but not statistically significant impacts, with effect sizes of 0.11 and 0.25, respectively. [Please insert Exhibit 5 here]

Conclusions:

Findings from this study demonstrate the success of the RAISE project in providing teachers with training and support at scale to help them change their instructional practices in order to foster metacognitive inquiry and support comprehension, particularly in science. These findings are consistent with positive findings from other studies of Reading Apprenticeship. The primarily positive, yet not statistically significant results for the full sample indicate that the study's sample size may not have been large enough to detect a modest size impact.

The results from this study point to several areas in need of further investigation. Specifically, the differences in impact by subject area and state need to be better understood. Further, SLI and the larger field would benefit from additional research on those factors that support bringing the model to scale and generating meaningful classroom level changes in instruction, particularly for ELA and history teachers. Overall, the study's findings demonstrate the potential of RAISE to address the paucity of content-specific reading instruction in U.S. secondary schools—even in science, where the need may be greatest.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

ACT. (2012). *Catching up to college and career readiness.* Retrieved from http://www.act.org/research/policymakers/pdf/CatchingUpToCCR.pdf

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., & Jones, B. (2011a). Integrating literacy and science in biology: Teaching and learning impacts of Reading Apprenticeship professional development. *American Educational Research Journal*, *48*(3), 647–717.

Greenleaf, C.L., Hanson, T., Herman, J., Litman, C., Rosen, R., Schneider, S., & Silver, D. (2011b). *A Study of the Efficacy of Reading Apprenticeship Professional Development for High School History and Science Teaching and Learning*. Final report to Institute for Education Sciences, National Center for Education Research, Teacher Quality/Reading and Writing, Grant # R305M050031.

Kemple, J., Corrin, W., Nelson, E., Salinger, T., Herrmann, S., & Drummond, K. (2008). *The Enhanced Reading Opportunities Study: Early Impact and Implementation Findings* (NCEE 2008-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

O'Reilly T., Weeks, J., Sabatini, J., Halderman, L., & Steinberg, J. (2014). Designing Reading Comprehension Assessments for Reading Interventions: How a Theoretically Motivated Assessment Can Serve as an Outcome Measure. *Educational Psychology Review, 26*(3), 403-424.

Somers, M. A., Corrin, W., Sepanik, S., Salinger, T., Levin, J., & Zmach, C. (2010). *The Enhanced Reading Opportunities Study Final Report: The Impact of Supplemental Literacy Courses for Struggling Ninth-Grade Readers*. NCEE 2010-4021. National Center for Education Evaluation and Regional Assistance.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1992–2013 Mathematics and Reading Assessments.

Appendix B. Tables and Figures

EXHIBIT 1. SELECTED BASELINE CHARACTERISTICS OF STUDY TEACHERS

	Total (N = 252)	Treatment (n = 130)	Control (n = 122)	*p* value
Gender and race				
Female	57.8%	56.3%	59.3%	0.674
Nonwhite	26.6%	26.8%	26.3%	0.958
Subject				
ELA	33.9%	34.2%	33.6%	0.930
U.S. history	33.5%	32.5%	34.5%	0.748
Biology	32.6%	33.3%	31.9%	0.815
Experience				
Average years teaching	9.94	10.35	9.54	0.649
Average years teaching the target subject	7.99	8.28	7.70	0.650
Percent certified as reading specialist	2.7%	3.6%	1.8%	0.403
Highest education level				
Bachelor's degree	37.2%	33.0%	41.2%	0.269
Master's degree	55.8%	60.7%	50.9%	0.167
Professional diploma or specialty	2.7%	3.6%	1.8%	0.675
Doctorate	0.9%	0.0%	1.8%	0.307
Other degree or credential	3.5%	2.7%	4.4%	0.489

Source. Study staff calculations based on the RAISE teacher surveys, year 1.

Note. The total teacher sample size is 252. Actual number of respondents for each subgroup varies because of missing data: 225–236 for the total sample, 113–116 for the treatment group, and 111–120 for the control group.

EXHIBIT 2. SELECTED CHARACTERISTICS OF STUDENT BASELINE SAMPLE

Student characteristic	Total (N = 14,747)	Treatment (n = 7,783)	Control (n = 6,964)	*p* value
Female	48.7%	49.5%	47.8%	0.146
Special education	11.3%	8.6%	14.1%	0.204
English language learner	13.8%	13.6%	14.0%	0.978
Eligible for free or reduced-price lunch	56.2%	56.7%	55.7%	0.846
Scored in bottom third on 8th grade state ELA or reading test	38.3%	39.9%	36.6%	0.266
Nonwhite	60.8%	63.4%	57.9%	0.636

Source. Study staff calculations on demographic data collected from study school districts.

Note. The total student sample size is 14,747, but demographic records were not reported for all students in the sample. Actual sample sizes vary by characteristic, depending on the completeness of the administrative data for each student in our sample.

EXHIBIT 3. IMPACT ESTIMATES FOR TEACHER SURVEY OUTCOMES, YEAR 2

Outcome measure (Teacher survey construct)	Adjusted treatment group mean	Adjusted control group mean	Difference (impact)	Standard error	*p* value	Effect size	N
1. Variety of text types	2.76	2.70	0.05	0.213	.798	0.04	206
2. Fostering student independence	4.36	3.34	1.02**	0.309	< .001	0.51	206
3. Traditional instructional strategies	4.35	4.16	0.18	0.315	.562	0.09	206
4. Teachers instructing metacognitive inquiry	0.76	0.84	- 0.07	0.114	.528	- 0.09	206
5. Teachers modeling metacognitive inquiry	0.96	0.88	0.08	0.096	.422	0.11	206
6. Students practicing metacognitive inquiry	1.91	1.49	0.42**	0.131	.001	0.46	206
7. Teachers instructing comprehension strategies	1.49	1.57	- 0.08	0.230	.719	- 0.06	206
8. Teachers modeling comprehension strategies	1.84	1.52	0.32	0.192	.096	0.23	206
9. Students practicing comprehension strategies	3.42	2.41	1.00**	0.238	< .001	0.62	206
10. Student collaboration	4.38	3.31	1.07**	0.405	.008	0.47	206

EXHIBIT 3. IMPACT ESTIMATES FOR TEACHER SURVEY OUTCOMES, YEAR 2

Outcome measure (Teacher survey construct)	Adjusted treatment group mean	Adjusted control group mean	Difference (impact)	Standard error	*p* value	Effect size	N
11. Student engagement	12.09	12.01	0.08	0.270	.760	0.05	206
12. Teachers' self-confidence in literacy instruction	39.31	36.67	2.63*	1.066	.014	0.41	206

Source. Study staff calculations on RAISE teacher surveys, year 2

* Significant at 5%

** Significant at 1%

EXHIBIT 4. IMPACT ESTIMATES FOR STUDENT OUTCOMES, YEAR 2

Construct	Adjusted treatment group mean	Adjusted control group mean	Difference (Impact)	Standard error	*p* value	Effect size	N
1.1: Participation in and contribution to class discussions	2.93	2.84	0.09	0.054	.102	0.12	11,398
1.2: Conferring	3.37	3.30	0.07	0.056	.194	0.09	11,400
2.1: Use of global reading strategies	2.94	2.92	0.03	0.039	.507	0.03	11,553
2.2: Use of problem-solving strategies	3.36	3.37	0.00	0.034	.916	0.00	11,534
2.3: Use of support reading strategies	2.59	2.53	0.05	0.051	.292	0.06	11,543
2.4: Integration of content and literacy activity	2.78	2.64	0.14**	0.052	.009	0.18	11,456
3.1: Metacognitive conversations	2.94	2.81	0.13**	0.045	.004	0.21	11,463
4.1: Reader identity	2.36	2.33	0.03	0.045	.533	0.04	11,419
5.1: Student identity	3.30	3.30	0.00	0.039	.999	0.00	11,426
6.1: Class time spent reading	2.73	2.62	0.10	0.079	.191	0.11	11,421
6.2: Variety of reading material	2.69	2.64	0.05	0.043	.238	0.05	11,396
6.3: Pages of reading per day	2.06	2.03	0.03	0.125	.782	0.03	10,831
7.1: Effort to learn	3.49	3.49	0.01	0.037	.865	0.01	11,390
7.2: Happiness and belonging	3.40	3.38	0.01	0.060	.835	0.01	11,386

EXHIBIT 4. IMPACT ESTIMATES FOR STUDENT OUTCOMES, YEAR 2

Construct	Adjusted treatment group mean	Adjusted control group mean	Difference (Impact)	Standard error	*p* value	Effect size	N
7.3: Engaging instruction	3.50	3.48	0.02	0.083	.834	0.02	11,401

Source. Study staff calculations on RAISE student survey, year 2

** Significant at 1%

EXHIBIT 5. IMPACT ESTIMATES FOR LITERACY ASSESSMENT SCORES BY STUDENT SUBGROUPS, YEAR 2

Group	Treatment group mean	Control group mean	Difference (impact)	Standard error	*p* value	Effect size	Improvement index	N
Full sample	- 0.02	- 0.16	0.13	0.100	.184	0.14	5.6%	10,173
California students	- 0.15	- 0.19	0.04	0.127	.765	0.04	1.6%	6,440
Pennsylvania students	0.16	- 0.08	0.24	0.175	.171	0.25	9.9%	3,733
Students in biology classes	0.01	- 0.29	0.31*	0.120	.010	0.32	12.6%	4,360
Students in ELA classes	- 0.04	- 0.24	0.20	0.135	.148	0.22	8.7%	2,936
Students in history classes	- 0.06	0.02	- 0.08	0.151	.618	- 0.08	-3.2%	3,449

Note. * significant at 5 percent ** significant at 1 percent

Source. Study staff calculations on ETS assessment data