

Abstract Title Page

Title: Measuring the Quantity and Quality of Instructional Interactions for Middle School Literacy Improvement

Authors: Sarah E. Carlson, Nancy J. Nelson-Walker, Patrick C. Kennedy, Gina Biancarosa, Jessica Turtura, Deanne A. Crone, Scott K. Baker, Hank Fien, & Jason Cohen

Abstract Body

Background & Context for the Study

As education policy in the U.S. focuses increasingly on the quality of classroom instruction, there is a growing need to identify valid constructs and develop reliable tools that measure student opportunities to learn academic knowledge and skills (Strong, 2011). Given the diversity of classroom inputs with the potential to impact student learning—e.g., content area, curriculum, resources, student autonomy, classroom management, assessment practices, teacher modeling, time on task, affective teacher-student relationships, and student practice opportunities—we face a series of important questions: How do these features of instruction relate to one another? How can they be measured? Which instructional practices should be emphasized for different purposes associated with student learning, and for which students?

Although there has been extensive research on the discreet features of classroom environments that contribute to student learning in elementary school classrooms, particularly in the context of early literacy instruction (e.g., Connor et al., 2009; Curby et al., 2011; Gersten et al., 2005; La Paro et al., 2011), these features have received less attention in middle grade literacy classrooms. Identifying the instructional practices that support adolescent literacy development is critical for several reasons. First, many middle school students do not read at a level required for success in higher education and the workplace (ACT, 2006; Kamil et al., 2008; NCES, 2011). Second, the reading demands placed on all students increase dramatically in the middle grades, as reading proficiency becomes the gateway to accessing content across all academic areas (Common Core State Standards, 2011). Third, efforts to synthesize effects of adolescent literacy research largely reveal that existing interventions are not effective for increasing student achievement outside of controlled studies (e.g., Enhanced Reading Opportunities and Striving Readers studies), indicating a need to provide intense interventions for struggling adolescent readers (e.g., Kamil, et al., 2008; Scammacca et al., 2007; Biancarosa & Snow, 2004). To identify practices that can increase adolescent literacy achievement, and in turn, design interventions and promote instruction that may be more effective in the middle grades, we contend it is important to study discreet features of adolescent literacy instruction implemented in middle school classrooms.

A common means of examining enacted instructional practices is direct observation. Standardized observation procedures are effective measurement tactics in instructional intervention research, and can improve our understanding of instructional and environmental factors that are causally related to student achievement (Kane & Staiger, 2012; Lipsey & Cordray, 2000). In this study we used two instruments to document features of adolescent literacy instruction and examine mediation effects between initial reading skill and student reading achievement: (a) the *Ratings of Classroom Management and Instructional Supports* (RCMIS: Doabler & Nelson-Walker, 2009), a measure of general instructional quality; and (b) the *Classroom Observations of Student-Teacher Interactions* (COSTI: Smolkowski & Gunn, 2012), a measure of contextual features of the classroom environment (e.g., number of minutes of instruction, content focus of instruction) and the rate of instructional interactions between teachers and students. These two instruments were used to examine instructional practices in Reading Intervention (RI) classrooms in schools participating in a longitudinal, multi-component project to evaluate middle school literacy instruction, student engagement, and data based decision-making.

Study Purpose & Research Questions

The purpose of this study is to examine the relative contributions of two different classroom observation instruments in an evaluation of classroom instruction in a comprehensive reading intervention program designed to support struggling readers. We explore whether instructional practices observed using each instrument mediate the effect of initial skill on student reading achievement, as a means for identifying practices that contribute to adolescent literacy development. Our specific research questions are: (a) Does general quality of instruction as measured by the RCMIS predict student reading achievement, above and beyond initial skill; (b) Does the rate of student-teacher interactions as measured by the COSTI predict student reading achievement, above and beyond initial skill; and (c) Does the RCMIS *or* the COSTI predict student reading achievement better than the other, after accounting for initial skill?

Setting & Participants

This study was conducted in 25 middle schools in six school districts in the Pacific Northwest. Trained staff observed 213 seventh grade Reading Intervention (RI) classrooms between one and four times across the 2011-2012 school year ($M = 2.36$). Altogether, 502 reading intervention observations were conducted in 2011-2012. Participants were 1,267 seventh grade intervention students and 124 RI instructors.

Description of the Intervention

The Middle School Intervention Project (MSIP) is a five-year (2010-2015), longitudinal study to evaluate a comprehensive intervention program for middle school students with reading difficulties in six districts in the Pacific Northwest. The intervention includes three components: (a) *Reading Intervention*, (b) *School Engagement Intervention*, and (c) *Data Based Decision Making*. Although the present study focuses only on the first component, it is important to note that these components, as a package, comprise a targeted, Tier 2 intervention that supplements established school-wide, universal Tier 1 reading and behavior support systems. Additionally, each of the intervention components are operationalized by participating districts; that is, MSIP is a partnership between researchers and participating districts, established to evaluate districts' existing instructional practices as they relate to providing comprehensive support to students who struggle in the area of reading.

Participating schools combine published intervention programs such as *Language!* (Greene, 2004), *Read 180* (Scholastic, 2006), and *Corrective Reading* (SRA, 1999) with other instructional approaches to deliver literacy interventions designed to (a) improve student reading outcomes on the state standardized reading assessment, and (b) enable students to meet reading demands in content-area classes (e.g., science, social studies). RIs are generally provided as a supplement to English Language Arts (ELA) instruction, and focus on critical reading content (e.g., word reading instruction, reading fluency, vocabulary, and comprehension strategy instruction) (Kamil et al., 2008; National Reading Panel, 2000; Pressley, 2000). RI classes are somewhat smaller than ELA or content-area classes, and meet for a minimum of 200 minutes per week. All RI instructors are trained in the reading programs used and in procedures of effective instruction for struggling students.

In addition, participating schools use test data (e.g., standardized reading assessments; progress monitoring measures) and other sources of evidence to determine which students receive RI and when they should exit the intervention. All students that receive RI also receive an engagement intervention (EI). Many schools implement formal engagement programs (e.g., Check In, Check Out; Check and Connect), while others use informal school-determined interventions that involve regular, personalized interaction with the same adult).

Research Design

A critical component of the MSIP project is that students are screened into the targeted Tier 2 intervention components based on 6th grade summative state assessment reading scores (Kamil et al., 2008; Scammacca et al., 2007). Schools are permitted to choose their own cut score, based on a rank order list of composite z-scores for each student. Composite scores are comprised of the average of (a) a z-score for student performance on a brief measure of oral reading fluency (using the sample mean score as the comparison), and (b) a z-score for student performance on the state reading assessment (using the state average score as the comparison mean). Each school identifies a cut score on the basis of available resources and the number of students they have the capacity to serve. All students who score below the school's identified cut score are screened into the comprehensive intervention and considered treatment students. Students who score above the school's selected cut score are considered comparison students.

To account for systemic variance associated with instructional practices and the assignment of students to classes (e.g., students may participate in multiple RI classrooms in a single day or across the year), we used a multi-level model nesting classrooms within students within schools by including weights for instructional variables of interest, individualized for each student based on the dosage of instruction they received from each classroom.

Measures. We used two measures to assess the diversity of classroom inputs (e.g., instructional content area, instructional structure, teacher modeling, teacher feedback, time on task, affective teacher-student relationships, student practice opportunities, classroom management, and the learning environment) and predict their impact on student learning.

RCMIS. The RCMIS is an 11-item Likert rating system targeting features of instructional quality (classroom management, delivery of instruction, and the learning environment). Observers use a 4-point scale to rate the overall quality of each item (Doabler & Nelson-Walker, 2009). Predictive validity coefficients of reading and mathematics outcomes range from .26 to .42. In the present study, the RCMIS was modified for use in middle school classrooms by replacing items targeting teacher-directed features of interactions with student-initiated approaches to learning and adding two items targeting student engagement (for a total of 13).

COSTI. The COSTI is designed to capture the duration, content domain(s), group structure(s), and rate of interactions between teachers and students (e.g., teacher models, feedback, student practice) during instruction. Studies indicate the COSTI can be used reliably to predict early literacy and math achievement (Doabler et al., in press; Fien, et al., 2012; Nelson-Walker et al., 2012; Smolkoski & Gunn, 2012). In our previous work with the COSTI, predictive validity coefficients with reading and mathematics outcomes have ranged from .25 to .55.

Student outcome measures. Student outcomes were assessed using scores from the state standardized reading achievement assessment and a measure of oral reading fluency (ORF). The state standardized reading assessment was an untimed, computer administered measure of student skill in comprehending text. The ORF measure was a one-minute, individually administered measure of student skill in accurately and fluently reading connected text.

Data Collection and Analysis

Observation procedures. Classroom observations using the COSTI and RCMIS were conducted in all RI classrooms in the fall, winter, and spring of the 2011-12 academic year. All observers (25 data collectors and 10 research staff) had experience participating in research studies and/or working in schools. Prior to each observation round, observers received training consisting of: (a) a review of content, (b) new content, (c) coding practice, and (d) reliability documentation. In total, observers received more than 24 hours of training across the year.

At the end of each training session, data collectors watched a video of classroom instruction and completed the RCMIS and COSTI. Observers' ratings were compared to the ratings of the observation coordinator and an inter-rater agreement index was calculated. In the fall, observers were also required to meet an inter-rater agreement criterion of .85 or better on all measures during a field checkout before observing independently in classrooms. Data collectors who did not meet criteria participated in additional training activities until they met criteria. Average field checkout reliability for the RCMIS was above .95 and for the COSTI was above .92. Reliability data were also collected for approximately 20% of all observations in each round to document maintenance of inter-rater agreement between trained observers. Efforts were made to pair observers (a) with each of the other observers assigned to their district(s), and (b) across regions for half of all reliability observations (i.e., cross-district observations). Average inter-rater agreement for the RCMIS was above .94 and for the COSTI was above .90.

Student assessment procedures. Student outcome measures were collected by district and project staff. Districts provided student scores on the state assessment in spring 2011 and spring 2012. Student performance on the state assessment in spring 2011 was used to create the z-score composite for intervention assignment and as a measure of initial reading skill, while student performance on the state assessment in spring 2012 served as the primary outcome measure. Trained staff collected ORF data in the fall of 2011 and spring of 2012. Spring 2011 assessment scores were used to create the z-score composite for intervention assignment.

Analysis procedures. HLM software (version 6.08; Raudenbush et al., 2004) was used to conduct a logical sequence of model testing proceeding from more-to-less constrained parameter estimates and to test hypotheses associated with each research question.

Study Results.

We found that the quantity of teacher instructional behaviors and student responses, as measured by the COSTI, did not predict student achievement gains on a standardized measure of reading comprehension. However, the effect of these behaviors varied significantly between schools. On the other hand, we found that overall instructional quality, as measured by RCMIS, significantly predicted student achievement across schools. In addition, we found that the RCMIS appears to be a better predictor than the COSTI when use in middle school RI classrooms. A complete correlation matrix and descriptive statistics are provided in Appendix B, Table 1. Results of the HLM analyses are reported in Appendix B, Table 2.

Conclusions

Initial findings from this study indicate that high quality instruction matters above and beyond the quantity of instruction students receive. Preliminary findings suggest the RCMIS can be used efficiently to document the quality of instruction provided to students in the middle grades. In addition, there appear to be unmeasured features of some intervention classes and schools that moderate the relationships between instructional factors and outcomes. Additional analyses are needed to explore relations between student-teacher interactions captured by the COSTI and student outcomes to determine whether the COSTI may be used to predict meaningful differences in middle grade instruction across RI classrooms.

Overall, these findings provide initial evidence for the types of observational tools that can be used to distinguish among the diversity of classroom inputs and how such inputs impact student learning. However, these measures may be limited in their ability to identify quality of instruction within a particular content area (e.g., reading comprehension). Future research is discussed in terms of developing new observational tools to identify specific instruction within the content area of reading comprehension and its impact on student learning.

Appendices

Appendix A. References

- ACT, Inc. (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: Author.
- Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report from the Carnegie Corporation of New York*. Washington DC: Alliance for Excellent Education.
- Common Core State Standards Initiative. (2010). Common core standards for English language arts & literacy in history/social studies, science, and technical subjects. Retrieved from Common Core State Standards Initiative: Preparing America's Students for College & Career website: http://www.corestandards.org/assets/CCSSI_ELA_Standards.pdf
- Connor, C. M., Morrison, F. J., Fishman, B. J., Ponitz, C. C., Glasney, S., Underwood, P. S., Piasta, S. B., Crowe, E. C., & Schatschneider, C. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, 38, 85-99.
- Curby, T. W., Stuhlman, M., Grimm, K., Mashburn, A., Chomat-Mooney, L., Downer, J., Hamre, B., & Pianta, R. C. (2011). Within-day variability in the quality of classroom interactions during third and fifth grade: Implications for children's experiences and conducting classroom observations. *Elementary School Journal*, 112(1), 16-37.
- Doabler, C., & Nelson-Walker, N. J. (2009). Ratings of Classroom Management and Instructional Support. Available from the Center on Teaching and Learning at the University of Oregon, Eugene, OR.
- Doabler, C. T., Fien, H., Nelson-Walker, N. J., & Baker, S. K. (in press). Evaluating three elementary mathematics programs for presence of eight research-based instructional design principles. *Learning Disability Quarterly*.
- Doabler, C. T., Baker, S. K., Smolkowski, K., Fien, H., Kosty, D., Miller, S. J., & Clarke, B. (2012). *Observing instructional interactions in kindergarten mathematics classrooms*. Manuscript in preparation.
- Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson-Walker, N.J., & Chaparro, E. A. (2012). An examination of the efficacy of a systemic intervention in early reading. *Exceptional Children*. Manuscript submitted for publication.
- Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting reading outcomes for first-grade English learners: An observational study. *Remedial and Special Education*, 26(4), 197-206.

- Greene, J.F. (2004). *LANGUAGE! A literacy intervention curriculum*. Longmont, CO: Sopris West.
- Kamil, M. L., Borman, G., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. K. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A Practice Guide*. (No. NCEE #2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. Retrieved from <http://ies.ed.gov/ncee/wwc>
- Kane, T. J. & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Policy and practice brief prepared for the Bill and Melinda Gates Foundation. Available from: <http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED529895>
- LaParo, K. M., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D., Early, D., Clifford, R., Barbarin, O., Carollee, H., & Burchinal, M. (2011). Quality in kindergarten classrooms: Observational evidence for the need to increase children's learning opportunities in early education classrooms. *Early Education and Development*, 20(4), 657-692.
- Lipsey, M. W. & Cordray, D. S. (2000). Evaluation methods for social intervention. *Annual Review of Psychology*, 46(3), 401-424.
- National Center for Education Statistics (2011). The nation's report card: Reading 2011. Retrieved December 6, 2011 from <http://nces.ed.gov/nationsreportcard/reading/>
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.
- Nelson-Walker, N. J., Fien, H., Kosty, D. B., Smolkowski, K., Smith, J. L. M., & Baker, S. (2012). Evaluating the effects of a systematic intervention on first grade teachers' explicit reading instruction. *Learning Disabilities Quarterly*. Manuscript submitted for publication.
- Nelson-Walker, N. J. (2010). Enhanced core reading instruction observation manual. Available from the Center on Teaching and Learning at the University of Oregon, Eugene, OR.
- Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 545-561). Mahwah, NJ: Erlbaum.
- Raudenbush, S. W., Bryk, A., Cheong, Y., & Congdon, R. T. (2004). HLM 6.08: Hierarchical linear and nonlinear modeling. Chicago, IL: Scientific Software International.

- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). *Interventions for adolescent struggling readers: A metaanalysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.
- Scholastic Research and Evaluation Department. (2006, July). *Compendium of READ 180 research* (Report 125916). New York: Scholastic.
- Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the classroom observations of student-teacher interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly*, 27(2), 316–328. doi: 10.1016/j.ecresq.2011.09.004
- SRA/McGraw-Hill. (1999). *Corrective reading*. TX: SRA/McGraw-Hill.
- Strong, M. (2011). *The highly qualified teacher: What is teacher quality and how do we measure it?* New York: Teachers College Press.

Appendix B. Tables and Figures

Table 1

Instructional Features Measured by the Ratings of Classroom Management and Instructional Supports (RCMIS) and Classroom Observations of Student-Teacher Interactions (COSTI): Correlations and Descriptive Statistics (N = 1267)

Variables	1	2	3	4	M	SD	Min	Max
1. Comprehension pretest	–				218.47	6.18	200	239
2. Comprehension post test	.596***	–			224.53	6.97	202	248
3. Instructional Quality	-.030	.018	–		42.23	5.85	0	51.33
4. Teacher Behaviors	-.205***	-.156***	.290***	–	292.18	188.26	0	1225.35
5. Student Behaviors	-.191***	-.153***	.284***	.971***	196.29	130.78	0	878.97

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2
Fixed Effects for Final School-Level Model

Estimate of fixed effects				
Variable	β	<i>SE</i>	<i>df</i>	<i>t</i>
Intercept β_0				
Intercept Υ_{00}	224.89	0.45	24	495.22***
Comprehension Slope β_1				
Intercept Υ_{10}	0.67	0.03	24	20.24***
Instructional Quality Slope β_2				
Intercept Υ_{20}	0.08	0.04	24	2.12*
Teacher Behaviors Slope β_3				
Intercept Υ_{30}	-0.001	0.004	24	-0.25
Student Behaviors Slope β_4				
Intercept Υ_{40}	-0.006	0.006	24	-1.11
Estimate of random effects				
Random effect	Residual variance	<i>SD</i>	<i>df</i>	Chi-square
	component			
Mean μ_{0j}	4.49	2.12	18	55.08***
Comprehension Slope μ_{1j}	0.01	0.10	18	31.88*
Instructional Quality Slope μ_{2j}	0.01	0.11	18	32.12*
Teacher Behaviors Slope μ_{3j}	0.0002	0.01	18	33.97*
Student Behaviors Slope μ_{4j}	0.0003	0.02	18	32.46*
Level 1 r_{ij}	25.44	5.04		

Note: Percent of variance explained demonstrates amount of variation attributed to between schools and within schools.

* $p < .05$, ** $p < .01$, *** $p < .001$