Teacher Looping Literature Search

Summary of Findings

Looping was first formally proposed as an educational strategy by Rudolf Steiner, philosopher and founder of Waldorf Schools in Germany, a hundred years ago. Looping is still practiced in Waldorf schools today (including the Waldorf School of Orange County –located in Costa Mesa). Deborah Meier, New York City Educator began using looping in 1974, introducing it to American public schools. Research first appeared in the 1990s. A few large scale studies continue to be conducted currently, with a resurgence in interest in recent years. The following are the researched benefits and concerns regarding the practice.

Description of the Practice of Looping

- Looping is when a teacher advances to a higher grade along with the students.
- The practice is largely used in elementary and middle schools.
- Waldorf schools have teachers loop from 1-8 grade.[7]
- There is evidence that Montessori schools and other private schools also utilize looping.[10]
- Documentation suggests public schools use looping for as little as 2 years, sometimes extending from 1-8 grade.[6,9,10,11,12]
- Looping during middle school years has been documented, as well.[6, 11.]
- No evidence of looping in High School was found.
- Studies mention that teachers are often either given a choice to loop or have teacher teams that loop with a student group.

Benefits of Looping

- Improved relationship between students and between teacher and students.[(all citations)]
- Many documented social-emotional gains, and parent involvement.[1,2,3]
- More efficient instruction: a gain of almost a month of teaching time in second year, since time for getting acquainted is eliminated and less review is needed.[1,2,3]
- Improved student discipline/classroom management.[1]
- Higher attendance rates. [5,12]
- Improved test scores particularly benefiting females and minorities.[9,11,12]
- Teacher gains such as opportunities for innovation were also mentioned.[2]

Concerns Around Looping

- The main concern is the possibility of a bad match between teachers and pupils, or among individual students or groups of students. [1,2,3]
- The possibility of having to put-up with a poor teacher for multiple years.

Finally, articles suggest that looping provides a structure to allow for certain positive outcomes, but not the cause of positive outcomes. Also studies suggest that looping thrives in the context of teacher buy-in and administrator support. **(Please note: References for footnotes found in attached Lit Review Matrix. Please see in attached Lit Review and articles for further information).**

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
1	2000(?)	AASA http://www.aasa.org/SchoolAdministratorArticle.aspx?id=14482	*In the Loop* Elementary and middle school mostly Overall Review Article	Provides brief history of looping and current thought. *"Looping provides a structure to allow for certain positive outcomes, but not the cause of positive outcomes."* Discusses National Middle School Association Study - "Florida Study" (A National Middle Schools Association study of looping at a school in Gainesville, Fla., by Paul George and colleagues) and Attleboro Mass. 1-8th grade study. Positive Student Outcomes include: • Improved relationship between students and between teacher and students (most important) • More efficient instruction • Higher attendance rates • Reduced student retentions • Fewer referrals of students to SPED • Improved student discipline/classroom management • Improved summer learning Positive Staff outcomes: • In Florida study, staff attendance improved Concerns/Obstacles: • Required effective teachers, and supports (from Admin, teachers union, parents) • Parent concerns of having an ineffective teacher for a long period • Willingness of teachers to move through the grades and return to original grade level

Foot note #	**Date**	**Authors/ source**	**Article Name/ School level /description**	**Summary**
2	2017	Teach thought Staff https://www.teachthought.com/learning/10-pros-and-cons-of-looping-in-education/	*10 Pros And Cons Of Looping In Education* Elementary Teacher Reflection	Reflections provided by a classroom teacher who has been with students for 3 years (total of 4 years with students by end of assignment) Positive: • Teacher-student relationship: o Shared experiences within the class o Perspective from past to future years o Students see teacher as trusted mentor o Developed a high level of caring and respect in classroom. • Strong, teacher-parent relationship • Understanding student needs o Teacher is aware of students strengths and weaknesses o Teacher aware of student triggers for good and bad o Easier monitoring progress for each student o Efficiency in teaching (no wasted days) • Promotes teacher innovation (with having to bring “new” experiences to students) • Benefits Classroom management o Established routines and procedures Concerns regarding Looping: • Teacher complacency due to getting too comfortable with students and families • Student miss opportunity to learn how to adapt to new situations • Students have less exposure to different teaching methods • If negative relationship exists, then persistence likely over the years • Difficult for a student to make new identity/patterns (i.e. “bad” student, labeled) • Teachers have to be comfortable and skilled with different grade levels *“Would I suggest looping to others? Yes, I would. Any teacher that is willing to put in the extra effort required to keep things fresh and innovative in their classroom can make this*

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
				model work wonders. There's no better sense of community that can exist than when it is created and allowed to exist over time."
3	1995	MAGnet Newsletter (Fall/Winter 1995, V4, No.1) http://web.archive.org/web/20010708041327/http://ericps.ed.uiuc.edu/eece/pubs/mag/magfal95.html	*Looping Through the Years:* *Teachers and Students Progressing Together* Elementary/1-8 Overall Review	Looping founded by Rudolf Steiner (father of Waldorf Education. A.K.A. Steiner education). Teachers with multi-year assignments with the same group of students have identified several social and academic advantages to looping. Social advantages: • reduced apprehension about the new school year and the new teacher after the first year (Hanson, 1995; Checkley,1995); • stronger benefits from the time spent on developing social skills and cooperative group strategies in the subsequent years (Hanson, 1995); • increased student self-confidence (Checkley, 1995) and a chance to overcome shyness (Mazzuchi & Brooks, 1992): • a stronger sense of community and of family among parents, students, and teachers (Checkley, 1995); and • greater support for children who look to school as a stabilizing influence in their lives. Academic benefits : • a gain of almost a month of teaching time, since time for getting acquainted is eliminated and less review is needed (Hanson, 1995; Mazzuchi & Brooks, 1992); • an increase in teacher knowledge about children's intellectual strengths and weaknesses in a way that is impossible in a single year; • an increase in the number of chances that are available to make connections during learning and over time (Zahorik & Dichanz, 1994); • More opportunities available to tailor the curriculum to individual student needs (Checkley, 1995).

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
				Concerns: • Possibility of a bad match between teachers and pupils, or among individual students or groups of students.
4	Last updated 2009	Cara Bafile Educational World https://www.educationworld.com/a_admin/admin/admin120.shtml	*In the Loop: Students and Teachers Progressing Together* Elementary/Middle Overall Review	Refers to MAGnet article above Benefits: • the promotion of stronger bonds between parents and teachers, teachers and students, and students and students; • greater support for children who need stabilizing influences in their lives; • a greater knowledge of students' strengths and weaknesses, allowing for increased opportunities for teachers to tailor curriculum to individual needs; • increased opportunities for shy students as well as others to develop self-confidence; • reduced anxiety about the new school year; and • a gain of almost a month of teaching time from the second year on, when the typical transition period at the beginning of the year is virtually unnecessary. Concerns: • an inappropriate match -- a personality conflict between student and student or teacher and student; • the possibility of having to put up with a poor teacher for multiple years; • the possibility, in this day of teacher mobility, that the teacher will not be there through the looping cycle; • less exposure to new students and teaching styles; • the difficulties faced by new students who enter a class that has looped; • the difficulty of adjusting to large school environments after being used to cloistered ones; and • the difficulty of separating at the end of the cycle, something that can be difficult for both teachers and students.

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
5	1997	https://www.brown.edu/academics/education-alliance/sites/brown.edu.academics.education-alliance/files/publications/looping.pdf	*Looping: Supporting Student Learning Through Long-Term Relationships* Brown university Pamphlet on Looping	Provides a summary of research from the 1990's and previously. Alludes to: Increases in ADA (92 % to 97.2%) – Mass. Sup. of Schools: Rappa, 1993 Drop-out rates at 1%
6	2005	Education World https://www.educationworld.com/a_issues/chat/chat037.shtml	*When Size Matters: Making Big Schools Feel Small (An Education World E-interview With Paul S. George)* An interview with Leading Middle School researcher (Paul George, University of Florida in Gainesville)	An interview with Paul George, a main researcher on Looping in Middle school. • Looping works well when the faculty and leadership are committed. • Looping -with teacher teams in middle schools is effective • Very few parents actually ask to have their children moved because of personality conflicts. • The "policy" that is needed is for the school leader to ensure that teaching talent is equalized (teams have equal strengths and weaknesses) and then not to cave in to parents who think their child will be better served on another team. • Possibly, Educators perceive greater benefit that parents and students
7	October 2015	https://phillywaldorf.com/multi-year-teaching-in-	An article from a Waldorf School on Teacher looping.	

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
		the-grade-school/		
8	March 2018	Hill, Andrew j, Jones, Daniel B. Economics of Education Review Volume 64, June 2018, Pages 1-12 https://www.sciencedirect.com/science/article/abs/pii/S0272775717306635 *See attached journal article	*A teacher who knows me: The academic benefits of repeat student-teacher matches* Elementary Journal Article (original research)	The researchers looked at students in grades 3 to 5 in North Carolina between 1997 and 2013 who were assigned to the same teacher two years straight. • Drawing on rich statewide administrative data, we observe small but significant test score gains for students assigned to the same teacher for a second time in a higher grade. • The effects are largest for minorities, • There is some evidence that gains are most evident for students with generally less effective teachers. (low performing teachers gain • We also provide suggestive evidence of spillover benefits: students assigned to classes in which a large share of classmates are in repeat student-teacher matches experience gains even if not previously assigned to that teacher themselves. This suggests that effects at least partly operate through improvements in the general classroom learning environment. From Educational Week.org: April 15, 2019 (https://blogs.edweek.org/edweek/teacherbeat/2018/03/looping_to_next_grade_boosts_learning_especially_kids_of_color.html • The estimated gain from students spending a second year with the same teacher ranges from .02 to .12 test score standard deviations. That's equivalent to an average student's score increasing by one or two percentiles, from say the 50th to 51st percentile. (The authors focused on math scores, but the appendix show that students also demonstrated an increase in end-of-year reading scores, though the gains were slightly smaller.)

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
9	2018	Findley, Maureen J. *ProQuest LLC*, D.E. Dissertation, The University of North Carolina at Chapel Hill The Impact of Looping in an Elementary School Setting https://eric.ed.gov/?q=looping&id=ED585827 (Dissertation)	*Impact of Looping in an Elementary School* Elementary dissertation (original research)	The purpose of this mixed methods study was to investigate the relationship between student assignments to a classroom practicing looping and student achievement on the state's End-of-Grade (EOG) exam in an elementary school setting. Findings: • The results of the quantitative portion of this study revealed that students' assignment to a classroom that practiced looping **did not have a statistically significant impact on overall student achievement nor did looping narrow the achievement gap**. • The results did indicate that assignment to a looping classroom **positively impacted mathematical achievement levels for students in the African American subgroup** but did not have a statistically significant impact for any other student subgroup. • The findings indicated that the teachers and administrators feel that looping is a positive experience for some students and had the potential to positively impact student achievement but is not a positive experience for all students. • Overall, the teachers and administrator found looping positively impacted relationships but had limited impact on achievement levels.
10	2017	Elena Nitecki https://files.eric.ed.gov/fulltext/EJ1142356.pdf	*Looping and Attachment in Early Childhood Education: How the Applications of Epigenetics Demand a Change*	Provides thorough history and overview of Looping. This qualitative study focuses on the relational benefits of looping, specifically in terms of attachment for young children, ages 0-5. Looping is especially important for young children, whose social emotional foundation is being built through attachments with parents and caregivers.

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
		Journal of the Scholarship of Teaching and Learning, Vol. 17, No. 2, April 2017, pp. 85-100. doi: 10.14434/josotl.v17i2.20840	Looping in Early childhood (0-5). A case study of a private pre-school using Montessori method. Journal Article (original research)	Two general research questions: "How does this preschool program use looping? What are the benefits and challenges associated with looping, as viewed through the lens of applied epigenetics?" Research itself is limited with just a one room school (13 kids) benefits were complex and intertwined Positive: • Teacher develops strong relationships with children over time and development • Consistency of relationships allow for open communication about stress at home • Behavior problems are minimal because teachers know what to expect from children • Extended learning time for older students who do not need transitional time at begging of year • Teachers have opportunities for extended leaning by having to be innovative over the years and adapt to different age groups • Parents develop bond with teacher over time. Leading to more investment in the school, volunteering and support of learning at home. Concerns: • Good fit between teacher and child or persistence of negative relationships over time. • The child's ability to transition and adapt to a new teacher/school after spending so long with a familiar teacher. • The teacher's ability to adapt to new age groups and the challenges each presents.

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
11	2010	Franz, Dana Pomykal; Thompson, Nicole L.; Fuller, Bob; Hare, R. Dwight; Miller, Nicole C.; Walker, Jacob School Science and Mathematics, v110 n6 p298-308 Oct 2010 https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1949-8594.2010.00038.x *See attached journal article	*Evaluating Mathematics Achievement of Middle School Students in a Looping Environment* Middle school Journal Article (original research)	This study was designed to determine if looping influenced middle school students' mathematical academic achievement. Student scores on the Mississippi Curriculum Test (MCT) were compared between sixth and eighth grade years for 69 students who looped during the seventh and eighth grades with a group of 137 students who did not loop. • Looping students (Females more than males) achieved statistically significantly greater growth on the MCT than their non-looping counterparts between sixth and eighth grades. • Greatest gains were made in 1st year of looping (second year of having same teacher) • Further, the data were disaggregated by gender, ethnicity, and socioeconomic status. • Findings indicate that looping may academically reengage students during the middle school years.

Foot note #	Date	Authors/ source	Article Name/ School level /description	Summary
12	2008	**Cistone, Peter**; Shneyderman, Aleksandr International Journal of Educational Policy, Research, and Practice: Reconceptualizing Childhood Studies, v5 n1 p47-61 Spr 2004. 15 pp. *See attached journal article	*Looping: An Empirical Evaluation* Elementary Journal Article (original research)	This evaluation study of the practice of looping (2 years) in a large urban school system was intended to explore its effect on student instructional outcomes, attendance, and retention rates, as well as to assess principals' and teachers' reactions to looping. Sample: (The looping sample included all those students from looping classes of selected elementary schools who were taught by the same teacher during the 1998- 1999 and 1999-2000 school years. This sample consisted of 612 students.) The results indicated that: • The Looping Sample outperformed their counterparts in the Matching Sample (in English reading, Mathematics). • Looping had a positive effect on student attendance. • Students in the Looping Sample had a significantly greater chance of being promoted to the next grade level. • Principals and teachers were in high agreement that looping had a positive effect on student learning in their schools. o Looping increases instructional time o Slower students have more time to learn the basic skills o Enhances the quality of the working relationship between teachers and students o Overall, enhances the effectiveness of classroom instruction. o In addition, although all teachers believed that they should be given a choice on whether to participate in looping, most teachers surveyed indicated that, given a choice, they would like to participate in looping again.

Economics of Education Review 64 (2018) 1–12

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

A teacher who knows me: The academic benefits of repeat student-teacher matches

Andrew J. Hill[a,*], Daniel B. Jones[b]

[a] Montana State University, Department of Agricultural Economics and Economics, P.O. Box 172920, Bozeman, MT 59717, USA
[b] University of South Carolina, Darla Moore School of Business, Department of Economics, 1014 Greene St., Columbia, SC 29208, USA

ARTICLE INFO

Keywords:
Student-teacher matches
Looping
Teacher specialization

ABSTRACT

We provide new empirical evidence that increased student-teacher familiarity improves academic achievement in elementary school. Drawing on rich statewide administrative data, we observe small but significant test score gains for students assigned to the same teacher for a second time in a higher grade. We control for selection into repeat student-teacher matches with teacher fixed effects and either student fixed effects or flexible controls for student past achievement. The effects are largest for minorities, and there is some evidence that gains are most evident for students with generally less effective teachers (as measured by value-added). We also provide suggestive evidence of spillover benefits: students assigned to classes in which a large share of classmates are in repeat student-teacher matches experience gains even if not previously assigned to that teacher themselves. This suggests that effects at least partly operate through improvements in the general classroom learning environment. Overall, our findings indicate that there may be potential low-cost gains from the policy of "looping" in which students and teachers progress through early school grades together, and may explain the recent experimental evidence that teacher specialization has negative effects on student achievement given that this likely decreases student-teacher familiarity.

1. Introduction

Among the many potentially important inputs into the education production function,[1] few have proven to be as consistently important as teachers. For instance, both Rockoff (2004) and Rivkin, Hanushek, and Kain (2005) find large gains from being assigned to a higher quality teacher, much larger, in fact, than gains from reductions in class size. Chetty et al. (2014a,b) show that being assigned to a higher quality teacher not only leads to higher test scores for students in grades 3 to 8, but also leads to better outcomes later in life. Understanding *how* teachers matter is clearly important, and the literature has explored several teacher-related factors relevant in education production (see, for example, Clotfelter, Ladd, & Vigdor, 2007). We propose a new mechanism that has not been directly explored, but is easily affected by policy: student-teacher familiarity from repeat student-teacher matches.

The potential importance of student-teacher interactions has been highlighted in previous research. Fryer (2016) finds that teacher specialization reduces student achievement, arguing that this may be caused by teachers having fewer interactions with each student. Several papers show that students who share demographic characteristics with their teachers perform better (Dee (2005), Hoffman and Oreopoulos (2009), Fairlie, Hoffmann, and Oreopoulos (2014), Ouazad (2014), Gershenson, Holt, and Papageorge (2016) and Egalite, Kisida, and Winters (2015)), and recent studies document that students show higher levels of achievement when their teachers are relatively optimistic about their future outcomes (Hill and Jones (2017); Papageorge, Gershenson, and Kang (2016)). Collectively, these papers suggest that familiarity with or understanding of students may play an important role in determining teachers' impacts on students.

Our paper addresses this possibility directly. Specifically, we provide evidence on the academic gains resulting from more thoroughly established student-teacher relationships in grades 3 to 5, early school grades in which teachers often provide emotional support in addition to instruction. We draw on administrative data covering the universe of public elementary school students and teachers in North Carolina. We identify students who – during the years we observe them – are assigned to the same teacher at least twice in different grades. We then estimate rich fixed effects models at the student-by-year level (incorporating either both student and teacher fixed effects or teacher

* Corresponding author.
E-mail addresses: andrew.hill6@montana.edu (A.J. Hill), daniel.jones@moore.sc.edu (D.B. Jones).

[1] See, for instance, Hanushek (2003) for an analysis of the efficacy of a variety of inputs in the education production function.

https://doi.org/10.1016/j.econedurev.2018.03.004
Received 19 October 2017; Received in revised form 13 March 2018; Accepted 16 March 2018
Available online 17 March 2018
0272-7757/ © 2018 Elsevier Ltd. All rights reserved.

fixed effects and flexible controls for past student achievement) to assess the impacts of being assigned to a teacher for the second time, taking performance on standardized end-of-grade tests as our outcome. We are identifying within-student growth in achievement in the second year that a student is assigned to the same teacher, and anticipate higher student test scores in the second year of the match compared to the first year if the establishment of student-teacher relationships impacts student achievement.

Ultimately, we find clear evidence that students perform better in the second year they are matched to a particular teacher. The estimated gain from a repeat student-teacher match ranges from 0.02 to 0.12 test score standard deviations. The lower end of that range stems from specifications with both student and teacher fixed effects, a rich model that removes a considerable amount of variation. Still, for the sake of comparison, this effect is similar in magnitude to the conditional black-white achievement gap in the same North Carolina administrative data.[2] Importantly, we rule out that results are driven by selection of students into repeat matches with teachers with whom they performed well in the past.

We present a series of results documenting heterogeneity in our main result. First, we find that the benefits of being rematched appear to be largest for students of generally less effective teachers (as measured by teacher value-added). And, second, we test how our result varies for different ethnic groups, finding that results are strongest for minority students. We think this is unsurprising given the average teacher is not from a minority group and therefore the initial "social distance" between a minority student and teacher may be relatively high, meaning there is more to gain from increased student-teacher familiarity.

The main results speak directly to the debate surrounding teacher specialization (often called "departmentalized instruction") in elementary education, a policy that is receiving increasing attention from policymakers and researchers. Although middle and high school teachers typically specialize in certain subjects (such as mathematics, English or social studies), elementary school teachers usually do not. Teacher specialization involves trading off the benefits of teachers sorting into subjects in which they have the comparative advantage (or perhaps that they simply prefer teaching) and the costs of reduced student-teacher familiarity from students rotating through teachers during the school day rather than having one teacher. The status quo suggests that the expected benefits of specialization outweigh the costs in high school, but do not do so in elementary school. Fryer (2016) conducts a field experiment in Texas elementary schools to test the productivity of teacher specialization, ultimately finding it has adverse effects on student achievement. Our findings of achievement gains from increased student-teacher familiarity support his claim that the costs of reducing student-teacher interactions may be sufficient to outweigh any benefits of teachers sorting into subjects.

Investigating repeat student-teacher matches also helps us understand some of the potential mechanisms through which the much-studied effects of student-teacher race congruence impact achievement. "Teacher like me" effects may operate through a variety of channels: the returns to general teacher effort may depend on the student-teacher race match if teachers use teaching styles that are more suited to students of the same race; teachers may explicitly or implicitly discriminate and allocate more student-specific effort to students of the same race; minority teachers may act as role models to minority students, showing minority students that there are returns to human capital investment; while another possibility is that teachers may have higher expectations of the students of the same race, and teacher expectations matter for student performance. Our finding of returns from student-teacher familiarity supports the presence of mechanisms like these: repeat student-teacher matches may improve the returns to general teacher effort through teaching styles that are both more adapted to the student and better understood by the student during the repeat match, they may increase student-specific teacher effort because of the more developed relationship between the student and the teacher; and teachers may see more potential in students when they know them well, leading to higher expectations. The role model channel, on the other hand, is likely more important in student-teacher race match effects than student-teacher repeat match effects.

While we study repeat student-teacher matches primarily as a window into the importance of student-teacher relationships, our results shed light on a policy which has received relatively little attention in the economics of education literature: "looping" . Looping, a policy in which whole classes (or most of the students within a class) are taught by the same teacher in sequential years, has been adopted in many countries (Cistone & Shneyderman, 2004). Cistone and Shneyderman (2004) conduct a survey of teachers in schools with looping and find that a large majority of teachers agree with the statement "Looping enhances the quality of the working relationships between teachers and students." [3]

Our results suggest that, at least with respect to gains in tests scores, looping is a worthwhile policy. While there is no official looping policy in our data, there are cases of sizable fractions of a class in some year t being rematched with the same teacher in year $t + 1$. We find that in such settings, the "unofficial" looping we observe has clear benefits not just for the students who were rematched, but also for other students in the class; repeat student-teacher matches appear to generate spillover benefits to other students in the class. Although our results on this are more suggestive than conclusive, this indicates that repeat student-teacher matches may improve the general classroom learning environment in addition to generating gains through more productive student-teacher interactions for rematched students.[4]

2. Empirical methodology

We estimate the effect of repeat student-teacher matches on academic achievement using the following model:

$$Y_{ijgt} = \theta RepMatch_{ijt} + f(Y_{i,t-1}, \alpha_i) + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt}$$

Y_{ijgt} is the end-of-grade test score of student i taught by teacher j in grade g during year t. It is standard normalized so that the magnitude of the estimated treatment effect is in units of standard deviations. The treatment variable $RepMatch_{ijt}$ indicates whether student i was also taught by teacher j in a previous grade and year (before grade g and year t). $RepMatch_{ijt}$ is zero in the first year of any match and one in the second year of a match. The function f(•) describes our control for student ability, which varies across specifications and is discussed in more detail below.[5]

[2] Hill and Jones (2016) estimate the size of the black-white achievement gap conditional on a variety of available controls, including lagged test scores, gender, race, neighborhood poverty status, and parental education.

[3] It is possible, of course, that there are long-run costs of looping or repeat student-teacher matches if they prevent students from learning how to adapt to different teachers. We think this is unlikely given that students will still be exposed to many teachers throughout their school careers even if they experience one or two "loops" or repeat matches, but cannot rule out this possibility in our analysis given the focus on short-term effects.

[4] In related research, Ly and Riegert (2014) find benefits from repeat classmates in high school. This suggests a potential mechanism operating in looping classes that is independent of the teacher match; students may gain from familiarity with peers even when the teacher changes.

[5] The set-up of the model considers treatment to be at the teacher-by-student level rather than the school level. As noted in the introduction, our results therefore speak indirectly to the policy of looping rather than being an explicit evaluation of a looping policy. In descriptive statistics not reported in the paper, we find that only ten percent of students are in a repeat match even when at least one other student in the same grade and school is in a repeat match. This tells us that the repeat matches we study in this paper are generally not implemented at a school or school-by-grade level, which prevents estimation of a meaningful school-level treatment effect in our study even though this would be of clear policy interest.

A student who is repeating a grade with the same teacher is not considered a repeat match. We do not want to consider these students as treated in our analysis because grade repetition has its own effect on school performance,[6] and we do not want to conflate our estimates of the effect of repeat student-teacher matches with this effect. Given that our analysis focuses on students in grades 3 to 5, $RepMatch_{ijt}$ can therefore only equal one for 4th and 5th graders. We include 3rd graders in the analysis even though they are never treated because 3rd grade scores improve our measure of student ability.

Students and teachers are not experimentally assigned to repeat student-teacher matches. It is therefore critical to control for factors that may affect the probability of students and teachers being involved in a repeat match in order to identify a causal effect of repeat matches. There are two first-order concerns.

First, the ability of students may be correlated with the likelihood of being involved in a repeat student-teacher match. For example, teachers who already teach different grades in different years may ensure that they are disproportionately assigned the better students when they teach a higher grade (who they can easily identify given they have recently taught them), or, alternatively, teachers may be more likely to teach a class in the subsequent grade when they have an above average class in the current grade. This would upwardly bias estimates of the effect of repeat student-teacher matches on school performance. On the other hand, administrators may try to reduce the burden of difficult or less able students by disproportionately assigning them to the same teacher in multiple years if there is a belief that difficult students are more easily managed when they are known by the teacher. We control for potential confounders related to student ability by flexibly controlling for the student's lagged test score $Y_{i,t-1}$ or including student fixed effects α_i in the model. Lagged test scores control for dynamic selection into repeat matches, but cannot capture unobserved factors, while student fixed effects deal with both observable and unobservable student characteristics that may be correlated with the likelihood of being involved in a student-teacher repeat match, but only those that are fixed over time. Ideally, we would control for both forms of selection, but including lagged independent variables as explanatory variables within a fixed effects framework may bias estimates towards zero. We therefore show that the main effect is robust to any combinations of these controls.

Second, teacher quality may be correlated with the probability of being in a repeat student-teacher match. For example, parents may be more likely to request the same teacher in multiple years when that teacher is of a high quality (and similarly more likely to request a new teacher when the previous teacher was of a low quality). Under this scenario, the estimated effect of repeat matches on academic achievement would also be biased upward. A related bias may arise from teachers sorting across schools. Higher quality teachers may disproportionately sort into schools in which it is more common for teachers to teach multiple grades if teachers view this as an opportunity to build longer-term relationships with students and these teachers are, on average, of a higher quality than other teachers. This would also upwardly bias the estimated effect. Teacher fixed effects β_j are included in the model to control for any selection into repeat student-teacher matches related to teacher quality or other teacher characteristics.

In a set of robustness checks, we also include teacher-by-grade and teacher-by-year fixed effects. These fixed effects deal with the possibility that some teachers may sort into grades in which they are better fits (in terms of student test scores) and their changing grades simultaneously generating repeat student-teacher matches. Using variation from teachers within a given grade ensures that we do not conflate gains from repeat matches with gains from teachers sorting into grades in which they are more productive.

Loosely speaking, the inclusion of student and teacher fixed effects means that the effect of repeat student-teacher matches – the coefficient on the $RepMatch_{ijt}$ indicator – is identified by comparing the end-of-grade test scores of the same student with the same teacher in a grade in which they have the same teacher for a second time to the grade in which they had the teacher for the first time.[7]

Finally, grade fixed effects δ_g and year fixed effects η_t control for any systematic variation in end-of-grade test scores across grades and time, as well any trends (over both time and grades) in the extent of looping observed in the data and academic achievement.

The model specified in Equation 1 allows us to explore potential selection by including an additional indicator in the model. $EverRepMatch_{ijt}$ is equal to one when students and teachers are in the first *and* repeating year of a repeat student-teacher match relationship. (To be clear, $EverRepMatch_{ijt}$ is always one when $RepMatch_{ijt}$ is one, but is also one in the first year that student i is taught by teacher j.) If the coefficient on $EverRepMatch_{ijt}$ is positive, for example, it means that higher-performing students are more likely to be involved in repeat student-teacher matches, which would be evidence of positive selection into repeat matches. We show in the results section how the coefficient on this variable changes with the inclusion of teacher and student fixed effects to better understand the nature of potential selection into repeat matches and to see whether our controls adequately deal with it.

In addition to the primary sources of bias related to student ability and teacher quality discussed above, an additional factor that may be correlated with student achievement and the likelihood of being involved in a repeat student-teacher match is school size. For example, it may be the case that students in smaller schools are more likely to be matched to the same teacher in multiple grades if teachers in smaller schools either teach more grades or have multi-grade classrooms. This could be particularly true in rural schools. If school size is also correlated with academic achievement (see Kuziemko (2006) and Gershenson and Langbein (2015), for example), our estimated effect of repeat student-teacher matches may be biased. School fixed effects will not fully capture school size effects given schools grow and shrink over time, especially considering our study sample of almost two decades. So, in addition to showing results are robust to school fixed effects, we therefore test whether results are sensitive to the inclusion of three additional time-varying fixed effects controls: the number of teachers who teach the same grade as teacher j in the school that year, class size, and an indicator for teachers teaching multiple grades in the given year (a proxy for a multi-grade classroom). The first of these is likely highly correlated with school size, but is in many ways a more direct control for potential bias caused by having fewer potential teachers with whom to match in smaller schools, while the second and third control for factors that may be indicative of particularly small schools or multi-grade classrooms. These specifications allow us to test whether any effect of repeat student-teacher matches persists not only when we are comparing the same student taught by the same teacher across two grades, but also a student who had the same number of potential other teachers in the school that year, and is in a class with the same number of classroom peers.[8]

In a final robustness check, we control for teacher experience to ensure that any gain is not driven by teachers in repeat matches necessarily having more experience than when they were first matched with the student. Specifically, given the evidence in Ost (2014) that

[6] See, for example, Eide and Showalter (2001), Jacob and Lefgren (2009), Manacorda (2012), and Hill (2014).

[7] It is worth noting for expositional purposes that it would be possible to include student-by-teacher fixed effects in the model – a dummy variable for *every* student-teacher combination – but, as would be expected, this model turns out to be over-fitted and soaks up most of the variation in the data, leading to less precise estimates of the effect.

[8] It is worth noting that we have estimated specifications where we restrict the sample to students more likely to eventually be engaged in repeat student-teacher match (based on a variety of observable characteristics of the student's school including urban-rural status of the school, number of students in the schools, etc.). Results are not reported but are very similar to our main results.

Table 1
Descriptive statistics: students and teachers in NC, 3rd to 5th grade, math, 1997–2013.

	Unique students		Unique teachers		Student-by-year observations	
	Mean	(s.d.)	Mean	(s.d.)	Mean	(s.d.)
Female	0.493	(0.500)	0.914	(0.280)		
Minority	0.403	(0.490)	0.193	(0.290)		
Black	0.273	(0.445)	0.129	(0.335)		
Ever repeat student-teacher match	0.030	(0.169)	0.094	(0.293)		
Teacher z-score			0.022	(0.785)		
Teacher years of experience			9.084	(9.238)		
Repeating student-teacher match					0.012	(0.111)
Number of teachers in grade (same school)					4.695	(2.136)
Class size					24.02	(11.11)
Indicator for "looping class": share of repeating student-teacher matches in class:						
> 0.2					0.022	(0.148)
> 0.4					0.012	(0.109)
> 0.6					0.007	(0.086)
> 0.8					0.003	(0.059)
Observations	2,111,082		70,471		5,135,946	

grade-specific experience matters for teacher productivity and that repeat student-teacher matches necessitate teachers switching grades and having different years of total experience and grade-specific experience, we show that results are not driven by changes in grade-specific teacher experience.

3. Data

We use data on the population of 3rd, 4th and 5th grade students and their teachers in public North Carolina elementary schools between 1997 and 2013 obtained from the North Carolina Education Research Data Center (NCERDC). Students are matched to the teacher who administered the end-of-grade test, which, for the early grades studied in this paper, is almost certainly the teacher who taught the class. We focus on end-of-grade mathematics test scores in this paper, showing in Appendix Table A1 that the effects of repeat student-teacher matches are similar, but smaller, for reading (which is consistent with the common finding that mathematics is more responsive to inputs in general).

Descriptive statistics are reported in Table 1. We have 2,111,082 unique 3rd to 5th grade students and 70,471 unique teachers who teach at least one 3rd to 5th grade class during the study period. As expected, about half the students are female. Forty percent of the students are minorities, the majority of whom are black. Turning to teachers, more than ninety percent are female, which is unsurprising given we are considering elementary schools. Notably, the share of teachers who are minorities is considerably smaller than the share of students who are minorities. This observation is particularly relevant when we relate our findings to the "teacher like me" literature; there is clearly a shortage of minority teachers if there is a desire to match students to teachers of the same race. On average, teachers have nine years of experience in the data.

Only three percent of students experience a repeat student-teacher match in the 4th or 5th grade,[9] and only fifteen percent of the 1,713 schools ever have a repeat match during the sample period. Although a small share, this represents over sixty thousand students given the large sample. Almost ten percent of teachers in the sample are ever involved in a repeat student-teacher match, which is higher than it is for students because typical students are exposed to only three teachers in over the three-year period from 3rd to 5th grade, while teachers are exposed to about 72 students during the same three-year period given an average class size of 24 students.

The final panel on the right describes the more than 5 million student-by-year observations used in the analysis. On average, the typical student has 4.7 potential teachers in a given grade at their school. The final four variables are indicators for "looping classes" – classes in which at least the given share of students in the class are in a repeat student-teacher match. We are interested in the extent to which repeat student-teacher matches are a product of traditional looping, the practice of entire classes of students progressing with their teachers to higher grades. In these types of classes, almost all the students in the class would be in repeat student-teacher matches. The descriptive statistics in Table 1 indicate that these classes are uncommon, and not the primary source of repeat student-teacher matches in our data. Only 0.3 percent of student-grade observations are in classes with a looping share of eighty percent, while 2 percent are in classes with a looping share of at least twenty percent.

4. Results

4.1. Main results

Our main results are reported in Table 2. The controls for past student achievement vary across the three panels, while other controls vary across the columns. We begin by discussing the model with no controls for lagged test scores in Panel A. Column 1 reports results from a baseline specification without controls, showing that 4th and 5th graders in repeat student-teacher matches score 0.123 standard deviations higher than other students on end-of-grade tests, on average. School fixed effects are included in Column 2 to control for the possibility that schools with teachers consistently teaching different grades in different years – perhaps schools with more administrative flexibility – are systematically better than other schools. We still observe the positive relationship between school performance and repeat student-teacher matches, indicating that school-specific factors do not drive the association between repeat matches and achievement.

Columns 3, 4 and 5 include teacher fixed effects, student fixed effects, and both teacher and student fixed effects, respectively. As discussed in the empirical methodology section, we consider these to be important controls for identifying the causal effect of repeat student-teacher matches on school performance.

Teacher fixed effects do not affect the magnitude of the estimated effect (0.129 in Column 3 in comparison to 0.123 in Column 1), suggesting that higher quality teachers are not disproportionately involved in repeat student-teacher matches. The inclusion of student fixed effects in Column 4, however, considerably attenuates the estimated effect.

[9] Recall that students matched to the same teacher while repeating a grade are not considered to be in a repeat match, so 3rd grade students can never be in a repeat match, but are included in the analysis for the estimation of the student fixed effects

Table 2
Effect of repeat student-teacher matches on standardized test scores in 3rd to 5th grade.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:						
Repeating student-teacher match	0.123***	0.149***	0.129***	0.036***	0.024***	0.021***
	(0.015)	(0.011)	(0.006)	(0.005)	(0.004)	(0.004)
Panel B:						
Repeating student-teacher match	0.039***	0.041***	0.021***	0.025***	0.023***	0.021***
	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)
Lagged test score: parametric (linear)	0.817***	0.794***	0.790***	−0.349***	−0.338***	−0.338***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Panel C:						
Repeating student-teacher match	0.033***	0.037***	0.025***	0.025***	0.020***	0.018***
	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)
Lagged test score: nonparametric (quintiles)	Y	Y	Y	Y	Y	Y
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,122,520	5,122,520	5,122,520	5,122,520	4,689,819	4,689,800

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The coefficient of interest falls to 0.036. The considerable decrease in the estimate suggests that students are positively selected into repeat student-teacher matches. Strikingly, though, after controlling for selection into repeat matches on student ability, students in repeat student-teacher matches still significantly outperform students in the first year of a match.

The specification in Column 5 includes both student and teacher fixed effects. Loosely speaking, we can interpret the results from this column as saying that the same student taught by the same teacher scores 0.024 standard deviations higher on end-of-grade tests when they are in a repeat student-teacher match. We consider this to be a small, but nontrivial, effect. It is of a similar order of magnitude to the within-school conditional black-white achievement gap (see Hill and Jones (2016), for example), and is subsequently shown to be considerably larger than the "teacher like me" effect, which has received considerable attention in the literature both in terms of estimating the direct effect of having a teacher of the same race on test scores and understanding the mechanisms through which the effect operates.

School fixed effects are included in the model in Column 6 for completeness. These are identified from the relatively infrequent circumstances in which students and teachers move schools. The estimated parameter of interest is very similar.

The models in Panels B and C of Table 2 add parametric and nonparametric controls for dynamic selection into repeat matches: a linear control for lagged test scores and lagged test score quintiles, respectively. The first observation from these panels is that the estimates across the six columns are considerably more stable than those in Panel A. This is because we are always controlling for some form of student ability. Most importantly, once we include teacher fixed effects in Column 3 (removing concerns about selection based on teacher quality), the estimated effects of interest in Panels B and C (0.021 and 0.025) are no longer statistically different from Panel A's preferred estimate that includes both teacher and student fixed effects (0.024 in Column 5).

There is evidence of mean reversion when we include both student fixed effects and lagged test scores. For example, in Columns 4, 5 and 6 of Panel B, a student with a higher test score (or, more specifically, a positive deviation from her mean test score) the previous year attains a lower test score this year. The potentially-biased estimates from these models (Columns 4, 5 and 6 of Panels B and C) are very similar to the estimates from the models that control for teacher-level and student-level selection in other ways. This indicates that if there is any bias arising from the inclusion of both student fixed effects and a lagged independent variable in our context, it appears to be negligible relative to the magnitude of the estimated effect.

Taking the results in Panels A, B and C as a whole, the estimated effect of interest is consistently within the 0.020 to 0.025 range as long as we include teacher fixed effects and some form of control for student ability. We focus on the model in Column 5 of Panel A for the remainder of the paper.[10]

Table 3 explores the extent to which student and teacher fixed effects deal with potential selection into repeat student-teacher matches. As discussed in the empirical methodology section, the coefficient on an additional variable $EverRepMatch_{ijt}$ – an indicator that is one for the first *and* repeating year of a student-teacher match – should be equal to zero in order for us to claim that assignment to repeat student-teacher matches is conditionally quasi-random.

The results in Column 1 of Table 3 include this additional variable, but exclude teacher and student fixed effects. This simple specification suggests positive selection into repeat matching. Students involved in repeat matches score 0.108 standard deviations higher on end-of-grade tests, on average, in both the first *and* repeating year of a match than other students. Interestingly, though, these students still experience an additional boost to achievement in the second year of a match of 0.017 standard deviations. The subsequent columns of Table 3 investigate the extent to which the key fixed effects controls attenuate the coefficient on this "ever repeat match" indicator, which we interpret as a measure of how much selection into repeat student-teacher matches our richer models capture.

We include both student and teacher fixed effects in Column 4. The coefficient on the "selection" variable is now zero. Repeat student-teacher matching boosts achievement by 0.022 standard deviations in this model, which is very similar to the corresponding estimate of 0.024 from Table 2, Panel A. Our takeaway from this table is that the positive selection into repeat matches is adequately accounted for when we use within-student, within-teacher variation to identify the effect.

We further probe the robustness of the finding in Appendix Table A2, showing that results are robust to including teacher-by-grade

[10] In Appendix Table A6, we report results from a model in which we add an indicator for being in a second repeat match with the teacher. To be clear, for a student who is matched to the same teacher in 3rd, 4th and 5th grade, this indicator is turned on in the 5th grade while the main indicator is turned on in both the 4th and 5th grade. Although we cannot draw strong conclusions from this specification given only two percent of repeat student-teacher matches are second repeat matches, we do find that students experience a precisely-estimated additional test score gain in the year of the second repeat match. The gains from student-teacher familiarity appear to grow as the familiarity increases.

Table 3
Potential selection in repeat student-teacher matches.

	(1)	(2)	(3)	(4)
Ever repeat student-teacher match with current teacher	0.108***	0.133***	0.017***	0.005
	(0.014)	(0.006)	(0.006)	(0.004)
Repeating student-teacher match	0.017***	0.012***	0.028***	0.022***
	(0.006)	(0.004)	(0.005)	(0.004)
Teacher fixed effects	N	Y	N	Y
Student fixed effects	N	N	Y	Y
Observations	4,689,819	4,689,819	4,689,819	4,689,819
R-squared	0.000	0.181	0.895	0.902

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

and teacher-by-year fixed effects. The estimates are attenuated, suggesting that part of the student achievement gain may be explained by teachers sorting into grades in which they are more productive combined with a simultaneous mechanical increase in their probability of being involved in a repeat match given they have changed grades. In addition, the smaller estimate from the specification with teacher-by-year fixed effects indicates that there may be positive classroom spillovers from repeat student-teacher matches. This is because the identifying variation in a model with teacher-by-year fixed effects comes from comparing repeat matchers and non-repeat matchers in the same class, so positive classroom spillovers would increase the test scores of students in the "control" group.[11] We also show in this table that results are robust to controlling for the number of teachers in the grade in the school, class size, a proxy indicator for multi-grade classrooms, and teacher experience. We show in Appendix Table A3 that the finding is not affected when we exclude grade repeaters from the analysis.

In Appendix Table A4, we report that the effect is largely driven by rematches in sequential grades. This table reports results from two models: one specification in which matches are only defined if they occur in sequential grades (Panel A), and another specification that includes an interaction between the original repeat match indicator and an indicator for being in a sequential grade (Panel B). This finding suggests that some of the benefits from increased student-teacher familiarity may arise from having better information about a student's past academic achievement, although it is difficult to disentangle this channel from simply having better information about the student on non-academic dimensions.[12]

Overall, the estimated effects of repeat student-teacher matches on student tests scores have a broad range: 0.02 to 0.12 standard deviations. Given the results discussed above, we consider the main sources of endogeneity to be student and teacher selection into repeat matches. We therefore interpret estimates at the lower end of this range in which we control for both of these to be the most accurate reflection of the true causal effect of repeat matches.

4.2. Heterogeneity by teacher quality

We now address a major concern related to repeat-student matches: what happens to students with low quality teachers who are then faced with low quality teachers for multiple years? Given the evidence that 3rd grade teachers have long-run effects on labor market outcomes (Chetty, Friedman, Hilger, Saez, Schanzenbach, & Yagan, 2011), parents may be justifiably concerned about repeat student-teacher matches having a particularly negative effect on student achievement when their children have low quality teachers. On the other hand, it may be the case that low quality teachers perform better when they know their students and have developed relationships with them. In this case, looping and repeat student-teacher matches may actually be a relatively low-cost tool to improve teacher performance for less effective teachers.

We probe heterogeneity in the effect of student-teacher matches by interacting the treatment variable with a proxy for teacher quality. In particular, we construct a simple measure of time-invariant teacher value-added by regressing student end-of-grade test scores on teacher fixed effects (and a variety of controls) using the sample of students not involved in repeat student-teacher matches.[13] We partition teachers into quartiles based on these teacher fixed effects, generating four teacher quality bins, and then fully interact our repeat-match indicator with the four teacher quality indicators. Note that it is important to control for the quality quartile of the student's teacher in the *previous* grade to ensure that we are not capturing the composite effects of previous teacher quality and current repeat student-teacher matching.[14]

Results are reported in Table 4. Column 1 of Table 4 shows that the general estimated effect of repeat matches is smaller, but still positive, in the reduced sample of 4th and 5th grade students for whom teacher value-added measures (VAMs) could be estimated for both their current and previous teachers. Turning to the interaction specification in Column 2, we find suggestive evidence of heterogeneity in the effect of student-teacher repeat matches by teacher quality. The gain from repeat matches for students with the highest quality teachers (the 4th VAM quartile) is 0.016 standard deviations *smaller* than the repeat match gain for students with the lowest quality teachers. Although this difference is imprecisely estimated, it suggests that repeat matches may be particularly useful for improving the performance of less effective teachers. Interestingly, the estimated gain from being in a repeat match with a teacher in lowest quality quartile (0.021 standard deviations) is similar in magnitude to the effect of replacing a student's teacher from the previous grade who was in the lowest quality quartile with a teacher who was in the second highest (0.019) or highest (0.024) quality quartile.

We also explore differential effects of repeat student-teacher matches along dimensions of teacher experience and teacher performance on credential exams. A teacher's z-score is a normalized measure of his or her performance on teacher licensing exams. Results in Column 3 of Table 4 show that the effects of repeat student-teacher matches are not affected by this measure of teacher ability, while the estimates in Column 4 show that teacher experience similarly has no influence on the gains from repeat student-teacher matches.

4.3. Heterogeneity by student race and student-teacher race congruence

Our analysis of the effects of repeat student-teacher matches on academic achievement has considerable parallels with the extensive

[11] We test directly for evidence of classroom spillovers in a later section.

[12] If there were teachers specializing in mathematics or reading in some years and not others, we would be able to probe this further by comparing the performance of students in repeat matches with teachers of the same subject and teachers of different subjects, but this is very uncommon in our data.

[13] There is a large literature exploring the validity of teacher value-added measures (Chetty, Friedman, & Rockoff, 2014; Rothstein, 2010). Given we are using our simple estimates as a somewhat coarse measure of teacher quality in this paper, we do not focus on these concerns here. As an aside, it is worth noting that the phenomenon of looping may generate its own concerns for generating valid measures of teacher value-added (such as those for tracking (Jackson, 2014), for example).

[14] The teacher VAMs we use to measure teacher quality are generated from the same data in which the student-teacher repeat matches we study occur. Given repeat matches are not commonly experienced by any teachers in our sample, this is likely to have a very small effect on our teacher quality measure. Saying that, it is important to note that interpreting VAMs as an exogenous teacher quality measure may be more problematic in many other settings, such as a study on the effects of looping involving several teachers who typically loop.

Table 4
Heterogeneity by teacher quality.

	(1)	(2)	(3)	(4)
Repeating student-teacher match	0.011**	0.021**	0.026***	0.033***
	(0.005)	(0.010)	(0.004)	(0.007)
Repeating student-teacher match X Teacher in 2nd VAM quartile		−0.013		
		(0.014)		
Repeating student-teacher match X Teacher in 3rd VAM quartile		−0.012		
		(0.014)		
Repeating student-teacher match X Teacher in 4th VAM quartile		−0.016		
		(0.014)		
Prev. teacher in 2nd VAM quartile		0.010***		
		(0.002)		
Prev. teacher in 3rd VAM quartile		0.019***		
		(0.002)		
Prev. teacher in 4th VAM quartile		0.024***		
		(0.002)		
Repeating student-teacher match X Teacher z-score			0.008	
			(0.005)	
Repeating student-teacher match X Teacher years of experience				−0.001
				(0.000)
Control for prev. teacher z-score	N	N	Y	N
Control for prev. teacher exp.	N	N	N	Y
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	2,190,826	2,190,826	3,350,348	3,352,679
R-squared	0.929	0.929	0.908	0.908

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

"teacher like me" literature.[15] This literature generally explores whether students perform better when they are exposed to teachers of the same race or gender. One of the channels through which student-teacher matches on observable characteristics may improve students' school performance is that teachers of the same race and gender may better understand the corresponding subset of students of that race or gender. Repeat student-teacher matches may allow teachers to develop a similar, if not better, understanding of their students than that provided by a race or gender match.

In addition, it may be the case that repeat student-teacher matches particularly benefit minority students. For example, deeper student-teacher relationships may make minority students feel more understood and included in the classroom. Furthermore, if minority students are more likely to come from more challenging family environments (such as single parent households), then teachers with better understandings of their specific backgrounds may be able to more adequately or appropriately address needs arising outside the classroom.

From a policy perspective, it is arguably much less costly to generate repeat student-teacher matches than student-teacher race matches. The prior requires teachers be certified to teach multiple grades – which is likely not too burdensome for the 3rd to 5th grade teachers studied in this paper, although, admittedly, may be more costly in terms of teacher training for higher grades – while the latter requires the composition of teachers to match the composition of students within schools – which may take both years of teacher training and targeted teacher sorting across schools. Promoting student-teacher race matches in the long run may also be undesirable given the clear potential for widening educational inequality and other adverse consequences.

In this subsection, we investigate the potential for repeat matches to have differential effects on minority students, and then we compare and contrast the repeat student-teacher matching effect with the "teacher like me" effect. Defining a student race indicator $Minority_i$, and the indicator $RaceMatch_{ijt}$ to be one when student i and teacher j are of the same race, the below specification is designed to achieve these goals.

$$Y_{ijgt} = \theta RepMatch_{ijt} + \pi(Minority_i \times RepMatch_{ijt}) + \rho RaceMatch_{ijt} + \sigma(Minority_i \times RaceMatch_{ijt}) + \alpha_i + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt}$$

For white students, the effect of repeat student-teacher matches on end-of-grade test scores is θ, while the effect of student-teacher race matches is ρ. For minority students, the corresponding effects are $(\theta + \pi)$ and $(\rho + \sigma)$. We report estimates for white students θ and minority students $(\theta + \pi)$ rather than estimates of differential effects π; the corresponding differential effects are reported in Appendix Table A5.

The estimates in Column 1 of Table 5 show that the effects of repeat student-teacher matches on end-of-grade test scores are positive for non-minority (white) and minority students, but, strikingly, the gains are significantly larger for minority students. We focus on the black-white difference in remaining columns so that estimates can be directly related to the literature on the black-white achievement gap, although the findings are very similar if we do not make this restriction.[16]

The model in Column 3 allows us to compare and contrast the effects of repeat student-teacher matches and the effects of same-race student-teacher matches. The boost to achievement from a repeat match is 0.028 standard deviations, while it is 0.007 standard deviations for a student-teacher race match (which is very similar to the corresponding student-teacher race match estimate in Egalite et al. (2015) using similar Florida administrative data). Both are statistically significant, but repeat matches are shown to be more beneficial. There is no additional gain from experiencing both a repeat student-teacher match and a same-race student-teacher match.

The results in Column 4 explore differences in these effects for white and black students. There are a few noteworthy observations: first, the repeat match effect is only present for white students when there is also a same-race match (white students gain 0.034 standard deviations on end-of-grade test scores in this circumstance), although this should be interpreted with the consideration that the majority of teachers in the sample are white; second, the repeat match effect is larger than the

[15] See, for example, Dee (2005), Fairlie, Hoffmann, and Oreopoulos (2014), Hoffman and Oreopoulos (2009), Ouazad (2014), Gershenson, Holt, and Papageorge (2016) and Egalite, Kisida, and Winters (2015).

[16] Restricting the sample to white and black students reduces the sample by about ten percent in North Carolina. In Column 2 of Table 6, we see that black students benefit more from repeat matches than white students, which just indicates that race differences in Column 1 are not driven by Hispanic students.

Table 5
Heterogeneity by student race.

	(1)	(2)	(3)	(4)
Effects on non-minority students:				
Repeating student-teacher match	0.017***			
	(0.004)			
Effects on minority students:				
Repeating student-teacher match	0.034***			
	(0.006)			
Effects on white students:				
Repeating student-teacher match		0.019***		−0.013
		(0.005)		(0.015)
Student-teacher race match				0.006**
				(0.003)
Repeating student-teacher race match X Student-teacher race match				0.034**
				(0.016)
Effects on black students:				
Repeating student-teacher match		0.034***		0.033***
		(0.006)		(0.007)
Student-teacher race match				0.012***
				(0.004)
Repeating student-teacher race match X Student-teacher race match				0.003
				(0.014)
Other effects:				
Repeating student-teacher match			0.028***	
			(0.006)	
Student-teacher race match			0.007***	
			(0.001)	
Repeating student-teacher match X Student-teacher race match			−0.007	
			(0.007)	
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,797	4,052,682	4,689,797	4,052,682
R-squared	0.902	0.903	0.902	0.903

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

"teacher like me" effect for both white and black students; and, third, the "teacher like me" effect is larger for black students than white students (0.012 in comparison to 0.006). The findings in Columns 1 to 4 show that repeat student-teacher matching is particularly beneficial to minority students and emphasize that the magnitude of the effect, although small, is nontrivial relative to other well-known effects associated with student-teacher matching.

4.4. "Looping" and probing the potential for spillovers resulting from repeat matches

Repeat student-teacher matches may improve school performance because teachers get to know their students and are able to adjust and target their teaching styles appropriately. Less directly, teachers' better understanding of their students may also help them provide discipline and improve the classroom atmosphere. It is therefore plausible that repeat student-teacher matches not only improve outcomes for the students involved in repeat matches, but also other members of the class.

Spillovers of this form are only distinguishable if not all students in a given class are in a repeat student-teacher match. If repeat matches are typically generated by the traditional practice of looping – entire classes progressing to the next grade with the teacher – the opportunity to estimate spillovers would be minimal. Perhaps surprisingly, the majority of repeat student-teacher matches in the data do not occur for entire classes; in fact, only five percent of repeat student-teacher matches in our data occur in classes where all other classmates are also looping.

Table 6
Class spillovers in effect of repeat student-teacher matches.

	(1)	(2)	(3)	(4)
Share x (see footnote)	*0.2*	*0.4*	*0.6*	*0.8*
Panel A:				
Looping class	0.015***	0.029***	0.030***	0.033***
	(0.004)	(0.005)	(0.006)	(0.009)
Panel B:				
Repeating student-teacher match	0.021***	0.011***	0.016***	0.020***
	(0.004)	(0.004)	(0.004)	(0.004)
Looping class	0.004	0.022***	0.018***	0.016*
	(0.004)	(0.006)	(0.007)	(0.009)
Panel C:				
Repeating student-teacher match	0.008	0.006	0.015***	0.020***
	(0.008)	(0.005)	(0.005)	(0.004)
Looping class	0.002	0.015**	0.016*	0.009
	(0.005)	(0.007)	(0.008)	(0.015)
Repeating student-teacher match X Looping class	0.016*	0.014*	0.003	0.008
	(0.009)	(0.008)	(0.010)	(0.016)
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,819	4,689,819	4,689,819	4,689,819

"Looping class" defined as class in which share of repeating student-teacher matches exceeds x, where x is stated in first row. All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We probe the potential for spillovers from repeat student-teacher matches in Table 6 by constructing an indicator for "looping classes" to include in the model. This variable is equal to one if the share of students who are involved in a repeat student-teacher match in the class exceeds some threshold share L. To be clear, this indicator will be switched on for any student who is currently in a class with at least the threshold share L of repeat matchers irrespective of whether the given student is in a repeat match themselves or not. Defining this variable as $LoopingClassL_{ijt}$, the below specification captures both the direct effect of being involved in a repeat student-teacher match and any potential spillovers from being in class with repeat matchers. We are now relying on student and teacher fixed effects to capture selection into both a repeat student-teacher match and a looping class.

$$Y_{ijgt} = \theta RepMatch_{ijt} + \mu_L LoopingClassL_{ijt} + \alpha_i + \beta_j + \delta_g + \eta_t + \varepsilon_{ijgt}$$

If both θ and μ_L are positive, there is evidence not only that students involved in repeat student-teacher matches show higher academic achievement, but also that all students in the class receive a performance boost when a sufficient share of their classmates are in repeat student-teacher matches.

Panel A of Table 6 reports achievement differences between students in looping classes and other classes. Column 1 indicates that students in classes in which at least 20 percent of their classmates are experiencing the teacher for the second time outperform students in other classes by 0.015 standard deviations, on average. Columns 2 to 4 show that this performance boost increases with the share of looping students in the class (40, 60 and 80 percent, respectively).

We cannot claim that this provides evidence of spillovers, of course, because the students actually in the repeat matches may experience all the gains. The model in Panel B – which also includes the original student-teacher repeat match indicator – allows there to be both a repeat match effect and a looping class effect. From Column 1, students in repeat student-teacher matches in classes in which at least 20 percent of their classmates are looping outperform other students by 0.021 standard deviations, but there is no indirect benefit from being in a class with other repeat matchers (an imprecisely estimated effect of 0.004 standard deviations). However, once the share of repeat matchers in the class reaches 40 percent (Column 2), there is both a direct gain from being in a repeat match and a spillover effect from being in a class with classmates who are in repeat matches. This gain persists when the threshold share for defining a looping class increases to 60 percent and 80 percent (Columns 3 and 4). The relative flatness of this estimate beyond the threshold share of 40 percent indicates that spillover effects do not appear to increase with the share of repeat matchers in the class once they are initiated.

Finally, Panel C of Table 6 allows potential spillover effects to differentially affect students who are in repeat matches and students who are not in repeat matches. Results reported in the first two columns of this panel suggest that achievement gains from repeat student-teacher matches only arise in classes with a sufficient share (20 or 40 percent) of other repeat matchers. There are no spillover effects for non-repeat matchers in these classes. Once 60 percent of students in a class are in repeat matches, though, Column 3 tells us that there is a direct benefit for repeat matchers (0.015 standard deviations), an indirect benefit for non-repeat matchers students (0.016 standard deviations), but no additional spillover effect if already in a repeat match (0.003 standard deviations). The results for those not in repeat matches are less clear when the share of repeat matchers reaches 80 percent of the class (Column 4), although there are now only a small number of non-repeat matchers in the class driving effects.

Overall, results in Table 6 are interpreted as suggestive evidence that repeat student-teacher matches generate benefits not only through specific student-teacher interactions, but also through changes in the general classroom learning environment, such as improvements in the classroom atmosphere or the enforcement of discipline.

5. Conclusion

In this paper, we assess the impact of repeat student-teacher matches in grades 3 to 5 on academic achievement. Drawing on administrative data from North Carolina, we estimate rich fixed effects models and find that students who are matched to a particular teacher for a second time score higher on standardized end-of-grade tests than they did in their first pairing with the same teacher: student-teacher familiarity improves student achievement.

These results shed light on the importance of student-teacher relationships in determining academic performance. We use repeat student-teacher matches as a window into the importance of teacher familiarity with students, but there are, of course, many other ways that teachers may have more established relationships or greater familiarity with certain students. For example, teacher specialization in elementary school is likely to *reduce* student-teacher familiarity, so the results in this paper serve as a caution for policymakers or school administrators implementing this increasingly-popular intervention. On the other hand, student-teacher demographic congruence (a “teacher like me”) may improve student-teacher relationships. While we certainly do not claim that our results can fully explain any of the benefits or costs of teacher specialization or having a “teacher like me”, the clear benefit of student-teacher familiarity we document in this paper may partly explain the mechanism through which these treatments impact student achievement. Another interpretation of our finding is simply that repeat matches allow teachers to reallocate time and effort away from getting to know their students to tasks that directly increase student learning, which, interpreted more generally, helps us think about how other policies that affect the within-classroom allocation of teacher time and effort may impact student performance.

Finally, our results also speak to a policy which has received little attention in the economics of education literature: “looping” . As noted in the introduction, looping is the practice of assigning entire classes (or most of the students from a class) to the same teacher for sequential grades. Education researchers have documented benefits of looping. For example, Cistone and Shneyderman (2004) compare average achievement of students in schools with looping to those without and find that students perform better in the looping schools. Our findings corroborate and build on their results; our data allow us to not only strip away more general school fixed effects, but also student and teacher fixed effects. We therefore identify a very clean estimate of the impact of repeat student-teacher matches. And, given the estimated effect is positive, we think that our paper motivates looping as a beneficial and relatively low-cost policy that should be given due consideration.

Acknowledgment

The authors thank Kara Bonneau of the North Carolina Education Research Data Center for assistance with the data.

Appendix

Table A1
Effect of repeat student-teacher matches on standardized language test scores.

	(1)	(2)	(3)	(4)	(5)	(6)
Repeating student-teacher match	0.085***	0.098***	0.104***	0.024***	0.013***	0.012***
	(0.013)	(0.009)	(0.005)	(0.004)	(0.003)	(0.003)
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,089,650	5,089,650	5,089,650	5,089,650	4,654,251	4,654,234
R-squared	0.000	0.106	0.153	0.885	0.880	0.881

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2
Effect of repeat student-teacher matches: robustness checks.

	(1)	(2)	(3)	(4)	(5)	(6)
Repeating student-teacher match	0.011***	0.008**	0.024***	0.022***	0.024***	0.029***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Indicator for teacher teaching multiple grades in current year					−0.035***	
					(0.005)	
Level of teacher-related fixed effects	Teacher-by-grade	Teacher-by-year	Teacher	Teacher	Teacher	Teacher
Student fixed effects	Y	Y	Y	Y	Y	Y
Number of teachers in grade fixed effects	N	N	Y	N	N	N
Class size fixed effects	N	N	N	Y	N	N
Teacher years of experience fixed effects	N	N	N	N	N	Y
Teacher grade-specific years of experience	N	N	N	N	N	Y
Observations	4,689,819	4,689,819	4,689,819	4,689,819	4,689,819	3,352,679
R-squared	0.905	0.915	0.902	0.902	0.902	0.908

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3
Effect of repeat student-teacher match excluding grade repeaters.

	Excluding students who are repeating a grade (1)	Excluding students who ever repeat a grade (2)
Repeating student-teacher match	0.022***	0.022***
	(0.004)	(0.004)
School fixed effects	N	N
Teacher fixed effects	Y	Y
Student fixed effects	Y	Y
Observations	4,545,898	4,449,383

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A4
Effect of repeat sequential-grade student-teacher matches.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:						
Repeating student-teacher match in sequential grade	0.124***	0.150***	0.133***	0.039***	0.027***	0.025***
	(0.016)	(0.012)	(0.006)	(0.005)	(0.004)	(0.004)
Panel B:						
Repeating student-teacher match	0.117***	0.139***	0.090***	0.008	−0.003	−0.007
	(0.024)	(0.020)	(0.014)	(0.012)	(0.010)	(0.010)
Repeating student-teacher match X Sequential grade	0.007	0.011	0.043***	0.031**	0.030***	0.032***
	(0.027)	(0.022)	(0.015)	(0.013)	(0.011)	(0.011)
School fixed effects	N	Y	N	N	N	Y
Teacher fixed effects	N	N	Y	N	Y	Y
Student fixed effects	N	N	N	Y	Y	Y
Observations	5,122,520	5,122,520	5,122,520	5,122,520	4,689,819	4,689,800

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A5
(alternate version of Table 5). Heterogeneity by student race.

	(1)	(2)	(3)	(4)
Repeating student-teacher match	0.017***	0.019***	0.028***	−0.013
	(0.004)	(0.005)	(0.006)	(0.015)
Repeating student-teacher match X Minority student	0.017***			
	(0.006)			
Repeating student-teacher match X Black student		0.015**		0.046***
		(0.007)		(0.016)
Student-teacher race match			0.007***	0.006**
			(0.001)	(0.003)
Repeating student-teacher match X Student-teacher race match			−0.007	0.034**
			(0.007)	(0.016)
Student-teacher race match X Black student				0.007
				(0.006)
Repeating student-teacher match X Student-teacher race match X Black student				−0.031
				(0.022)
Repeating student-teacher match X Female student				
Teacher fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Observations	4,689,797	4,052,682	4,689,797	4,052,682
R-squared	0.902	0.903	0.902	0.903

All regressions include year and grade fixed effects. The dependent variable is the standardized test score from end-of-grade mathematics tests. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A6
Effect of second repeat student-teacher match on standardized test scores in 3rd to 5th grade.

	(1)
Repeating student-teacher match	0.023***
	(0.004)
Second repeat match	0.057***
	(0.027)
School fixed effects	N
Teacher fixed effects	Y
Student fixed effects	Y
Observations	4,689,819

All regressions include year and grade fixed effects. Robust standard errors clustered at school level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2018.03.004.

References

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from project star. *Quarterly Journal of Economics, 126*, 1593–1660.

Chetty, Raj, Friedman, JohnN., & Rockoff, JonahE. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review, 104*, 2593–2632.

Cistone, Peter, & Shneyderman, Aleksandr (2004). Looping: An Empirical Evaluation. *International Journal of Educational Policy, Research, and Practice: Reconceptualizing Childhood Studies, 5*, 47–61.

Clotfelter, CharlesT., Ladd, HelenF., & Vigdor, JacobL. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673–682.

Dee, ThomasS. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review, 95*, 158–165.

Egalite, AnnaJ., Kisida, Brian, & Winters, MarcusA. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review, 4*, 44–52.

Eide, EricR., & Showalter, MarkH. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review, 20*, 563–576.

Fairlie, RobertW., Hoffmann, Florian, & Oreopoulos, Philip (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review, 104*, 2567–2591.

Fryer, Roland (2016). The 'pupil' factory: specialization and the production of human capital in schools. NBER Working Paper w22205.

Gershenson, Seth, & Langbein, Laura (2015). The effect of primary school size on academic achievement. *Educational Evaluation and Policy Analysis, 37*, 135S–155S.

Gershenson, Seth, Holt, StephenB., & Papageorge, NicholasW. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review, 5*, 209–224.

Hanushek, EricA. (2003). The failure of input-based schooling policies. *Economic Journal, 113*(48), F64–F98.

Hill, AndrewJ. (2014). The costs of failure: Negative externalities in high school course repetition. *Economics of Education Review, 4*, 91–105.

Hill, Andrew, & Jones, Daniel (2017). Self-fulfilling prophecies in the classroom. Available at SSRN: https://ssrn.com/abstract=3016029.

Hill, Andrew J., & Jones, Daniel B. (2016). "Paying for whose performance? Teacher incentive pay and the black-white gap." Working Paper.

Hoffmann, Florian, & Oreopoulos, Philip (2009). A professor like me: The influence of instructor gender on college achievement. *Journal of Human Resources, 44*(2), 479–494.

Jackson, C. Kirabo (2014). Teacher quality at the high-school level: the importance of accounting for tracks. *Journal of Labor Economics, 32*(4), 645–684.

Jacob, BrianA., & Lefgren, Lars (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics, 1*, 33–58.

Kuziemko, Ilyana (2006). Using shocks to school enrollment to estimate the effect of school size on student achievement. *Economics of Education Review, 25*, 63–75.

Ly, Son Thierry, & Riegert, Arnaud (2014). "Persistent classmates: how familiarity with peers protects from disruptive school transitions." PSE Working Papers n2013-21.

Manacorda, Marco (2012). The cost of grade retention. *Review of Economics and Statistics, 94*, 596–606.

Ouazad, Amine (2014). Assessed by a teacher like me: Race and teacher assessments Economics of Education Review. *Education Finance and Policy, 9*(3), 334–372.

Ost, Ben (2014). How do teachers improve? The relative importance of specific and general human capital. *American Economic Journal: Applied Economics, 6*(2), 127–151.

Papageorge, Nicholas W., Gershenson, Seth, & Kang, Kyungmin (2016). "Teacher Expectations Matter." *No. 10165. Institute for the Study of Labor (IZA)* Working Paper.

Rivkin, StevenG., Hanushek, EricA., & Kain, JohnF. (2005). Teachers, schools and academic achievement. *Econometrica, 73*(2), 417–458.

Rockoff, Jonah (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review, 94*(2), 247–252.

Rothstein, Jesse (2010). Teacher quality in educational production: tracking, decay, and student achievement. *Quarterly Journal of Economics, 125*, 175–214.

Evaluating Mathematics Achievement of Middle School Students in a Looping Environment

Dana Pomykal Franz
Nicole L. Thompson
Mississippi State University

Bob Fuller
Starkville School District

R. Dwight Hare
Nicole C. Miller
Mississippi State University

Jacob Walker
Mississippi State University

Looping, a school structure where students remain with one group of teachers for two or more school years, is used by middle schools to meet the diverse needs of young adolescents. However, little research exists on how looping effects the academic performance of students. This study was designed to determine if looping influenced middle school students' mathematical academic achievement. Student scores on the Mississippi Curriculum Test (MCT) were compared between sixth and eighth grade years for 69 students who looped during the seventh and eighth grades with a group of 137 students who did not loop. Looping students achieved statistically significantly greater growth on the MCT than their nonlooping counterparts between sixth and eighth grades. Further, the data were disaggregated by gender, ethnicity, and socioeconomic status. Findings indicate that looping may academically reengage students during the middle school years. Advantages and disadvantages of looping at the middle grades are discussed.

Middle school mathematics could be referred to as *math in the middle*. For example, students are ability-tracked, often for the first time. Mathematical concepts become increasingly abstract and complex, demanding a thorough understanding of mathematics learned in earlier grades (Carraher & Schliemann, 2007; Kieran, 2007). Teaching styles and classroom and school day organization begin to vary. Teachers may be secondary-trained content specialists, middle-level specialists with a mathematical content emphasis, or elementary educators with varying levels of mathematical content knowledge. Further, teachers perpetuate the beliefs that males are simply better at mathematics than females as seen in teacher behavior (Good & Brophy, 2007), despite evidence from Trends in International Math and Science Study, TIMSS, (National Center for Education Statistics, 2003) along with other research (Fennema & Leder, 1998; Hyde, 1990; Noddings, 1993), which indicates there are no significant differences between males' and females' abilities in mathematics and/or sciences. Young adolescents need well-prepared teachers in middle schools who can bridge the transitions from elementary school mathematics to increasingly abstract middle school courses, break perceived barriers to access mathematics, and keep students motivated and excited about learning mathematics.

Recent research suggests motivational factors such as students' intellectual support by teachers and peers are extremely important in the achievement of mathematics (National Mathematics Advisory Panel, 2008; Shores & Shannon, 2007). Achievement gaps for minority students in mathematics, while trending downward, is still greater than 30 points between African American and Caucasian students and slightly less than 30 points between Hispanic and Caucasian students in eighth grade (National Center for Education Statistics, 2008). When minority status is combined with poverty, the need for a supportive learning environment is essential (Payne, 2005). Young adolescents need organizational structures in middle schools that may reduce the effect of poverty on academic achievement. Research has shown that the greater the poverty level of a school, as indicated by the number of students receiving free or reduced lunch, the lower the overall academic achievement (Brown, Roney, & Anfara, 2003; Mertens & Flowers, 2003).

Middle-level education is challenging for both teachers and students. Students are deciding what their attitudes are about mathematics, education, and even careers. Increased student achievement, participation and attendance in school, and reduced disciplinary issues are all attributed to the development of a long-term teacher-student relationship (George & Lounsbury, 2000). Looping, as an organizational structure, supports the development of this type of relationship. Further, looping may provide an environment that fosters student achievement in middle school mathematics.

Background

At the time of its inception, the idea behind the middle school was to create a school for older children who were *in the middle* (Alexander, 1995) that needed a developmentally appropriate and responsive learning environment. The students who were *in the middle* were not quite ready for the challenging aspects of high school but were definitely able to work more independently than elementary school students were. The resulting middle school was designed to feature several educational components that foster all areas of young adolescent development: physical, intellectual, emotional, social, and moral. The most common elements of the middle school concept that best meet the needs of young adolescents are interdisciplinary teams, advisory, transition and exploratory programs, and teaching a curriculum that employs varied instructional practices and is “relevant, challenging, exploratory, and integrated” (National Middle School Association, 2003). Teacher teams with common planning time that meet regularly to discuss the needs of their students are also more responsive to students’ academic and affective needs (George & Alexander, 2003; Kellough & Kellough, 2003; Powell, 2005). Although several decades have passed since this original vision was unveiled, the ideas for what middle school should be have not changed.

In addition to the academic emphasis of schooling and the pressures of standardized assessments, middle school teachers and administrators have to deal with the ever-changing needs of adolescents. Young adolescents are experiencing a vast amount of change at this time in their lives. Physical and intellectual development occurs concurrently with each greatly influencing the other (Powell, 2005). How a young adolescent feels physically affects every other aspect of his or her life. Added to these basic physiological changes is the need to fit-in and the desire to belong—the young adolescent lives in a tumultuous world that is constantly evolving. Young adolescents need teachers who understand their development and are knowledgeable about the world in which they live (Jackson & Davis, 2000). This critical time in their lives emphasizes the need for stability and structure. Young adolescents require a learning environment that supports them so they are able to be successful in school. Looping is a structure that allows teachers to blend increasingly complex content and specialized subject matter with a nurturing learning environment (Kellough & Kellough, 2003; Powell).

Advantages of Looping

Looping is an organizational structure that keeps students and teachers together for two or more school years. Comer (2001) posited that significant relationships must be established between teachers and students for significant learning to occur. The looping environment creates an atmosphere where young adolescents are able to take risks with learning because significant relationships are developed between the teacher and students. This established relationship enables students to achieve success in a variety of areas. As noted by Jackson and Davis (2000), every young adolescent needs the opportunity to experience success in school. Looping is a structure that provides this opportunity. Although most of the reported benefits of looping do not directly address academic achievement, the context for improved academic success is created. Likewise, students in mathematics classrooms need to be encouraged to take risks. Increasing cognitive demands in mathematics are dependent on students conjecturing about their work and making connections to concepts previously learned. Being in a classroom that supports and even encourages risk-taking is imperative for mathematical growth (NCTM, 2000).

Looping provides teachers time to get to know their students’ strengths and weaknesses (Crosby, 1998; Elliot, 1998; Forsten, Grant, Johnson & Richardson, 1997), track students’ long-term progress (Forsten, Grant & Richardson, 1999), and improve instructional planning (George & Lounsbury, 2000; George & Shewey, 1997; Manning & Bucher, 2001). The extended time also enhances the teacher’s role as a “guide and facilitator of learning” (Kasak, 2004, p. 243). Long-term relationships facilitate the development of effective instructional approaches that meet students' specific needs (George & Shewey, 1997; McLaughlin & Doda, 1997). This environment is needed for students as their mathematical understandings develop. Sleep and Ball (2007), in their research on mathematics' teachers, described the impreciseness of students' discourse about mathematics. They further emphasized the need for teachers that can listen and interpret their students' mathematical statements. Teachers who have long-term relationships with students develop insights into the ways their students think and talk, helping interpret classroom discourse.

Looping-developed relationships lead to continuity, generating better learning environments and positive relationships between teachers, students, and parents. Jackson and Davis (2000) posit that time provided by looping allows students and parents to develop healthy attachments to teachers. The mathematical field gains a decided benefit. Specifically, females and underrepresented groups need support as they develop their mathematical interests. Despite the efforts of national funding agencies to narrow the gap between males' and females' interests in mathematics and related STEM fields, females still lag behind males. Since career exploration begins in middle school, having environments that support female or minority interests is important. Positive teacher-student relationships enhance this process. Numerous studies support this idea (e.g., American Association of University Women, 2004; Darke, Clewell, & Sevo, 2002; Dentith, 2008; McCullough, 2003).

A great deal of time is spent each school year in the "getting to know you" process. Time normally dedicated to this relationship building is reduced when looping. George and Lounsbury (2000) and George and Shewey (1997) studied long-term teacher–student and teacher–parent relationships. They found that all looping participants agreed a greater sense of community developed as time progressed. Improved relationships between teachers, students, and parents lead to increased satisfaction in the educational experience (Forsten et al., 1999). Each group did not have to experience normal adjustment periods that occur every school year. Further, time is saved by already understanding students' knowledge levels (Crosby, 1998; Darling-Hammond, 1997); this is particularly helpful for students facing academic or social challenges (Forsten et al., 1997). With increasingly complex and abstract concepts being presented in middle-level mathematics courses, the gained instructional time provided by looping gives teachers additional opportunities to assess their students' skills and understandings.

Disadvantages of Looping

While the advantages of looping are notable, disadvantages do exist. The primary concern associated with looping is the teacher. Veteran teachers, who have more experience with classroom management and varied instructional strategies, tend to experience greater success when looping than beginning teachers do (Simel, 1998). Seasoned teachers have larger "bags of tricks" and tend not to run out of new and innovative ideas for teaching, building rapport with students and parents, and managing the classroom. Further, all teachers and administrators associated with looping need to be educated about the structure and complexity of it. Proper training and implementation is necessary so that looping teachers and administrators understand all of the components of looping and are aware of the investment that will be made. Additionally, administrators need to understand the intricacies of looping and realize it does create a positive learning environment, but looping is not necessarily a structure that will meet the learning needs of students with disabilities, second language learners, or other students with additional social or academic needs (Forsten et al., 1997; McAteer, 2001; Simel).

Parental concerns about looping exist as well. Parents of looping students tend to worry that their child will be taught by an ineffective teacher for two or more years and may feel that their child is "stuck." In addition, parents, students, and teachers might experience difficulty cultivating the long-term relationship associated with looping which might reduce its positive impact (Nichols & Nichols, 2002; Simel, 1998). Finally, students who transition into and out of the looping cycle might experience feelings of being an "outcast" either in the looping classroom or in their new school (Simel). Looping creates cohesiveness among all people involved and being a member of a tightly knit group may be difficult. Likewise, leaving such a group can have a profound impact on a student. Therefore, looping works best when students begin and end the loop together; yet this does not always occur given the transient nature of society. Everyone involved in the looping cycle needs to communicate effectively with the other, as relationship building is the backbone of the looping process.

Purpose of the Research

The majority of research about looping tends to focus on the elementary school and be qualitative in nature. Therefore, the purpose of this research was twofold. The primary objective was to determine if middle grades students who looped performed better on the mathematics portion of the Mississippi Curriculum Test (MCT) than did their nonlooping peers. Expanding on that notion, the researchers were interested in finding connections between gender, ethnic group, and socioeconomic status and looping, and if those connections impacted MCT performance. The secondary reason to conduct the research was to contribute to the limited research base regarding middle grades looping.

Methods

Setting

The Middle School. This research occurred in a middle school located in rural east central Mississippi. The middle school is the only school in the district serving grades seven and eight. Sixth grade students are housed at a separate stand-alone campus. In the mid-1990s, both the sixth grade school and the middle school implemented elements of the middle school philosophy, including interdisciplinary teaming (i.e., mathematics, language arts, science, and social studies teachers collaboratively working with the same group of students), flexible scheduling (i.e., interdisciplinary teams' ability to manipulate their daily schedule to meet instructional goals), (George & Alexander, 2003), and common planning time (i.e., a specified time set aside each day for interdisciplinary teams to meet and discuss ideas, issues, and student needs) (Mertens, Roney, Anfara, & Caskey, 2007). The district has consistently scored near or above the state average on the statewide criterion referenced MCT.

The average enrollment at the middle school is 650 students. The teacher to student ratio is approximately 1:30. About 65% of the district's students qualify for free or reduced lunch. The district has a highly transient population due to both its proximity to a major university and the mobility of its low socioeconomic population.

School Context. All students were placed on interdisciplinary teams. To maintain equity and stability within core teams (i.e., language arts, science, social studies, and mathematics), school administration assigned teachers to teams so that each team had equal ratios of experienced and less-experienced teachers. Time was provided to teachers for individual planning as well as for common planning time. To ensure the same content was taught in all classrooms, members of the faculty met bi-monthly within subject area departments to focus on curriculum maps. Curriculum maps are living documents which are teacher-created, used to guide content and assessments throughout the school year (Hayes Jacobs, 2004). These maps were constantly revised to meet the academic needs of students. Common assessments were also given to students each nine-week grading period, ensuring that common instruction occurred.

Sample

This study tracked the mathematics achievement of 206 students during their seventh and eighth grade years. The students' sixth grade mathematics MCT scores were used as baseline data. The treatment group (looping team) consisted of 69 students who looped (i.e., remained with the same team of teachers) during only grades 7 and 8; the comparison group (nonlooping teams) consisted of 137 students who did not loop during grades 7 and 8. The looping team consisted of 35 females and 34 males. The nonlooping teams were composed of 82 females and 55 males. Overall, there were 124 African American students, 78 Caucasians and 4 students of "other" races. The looping team consisted of 36 African Americans, 31 Caucasians and 2 other race students. The nonlooping teams consisted of 88 African Americans, 47 Caucasians and 2 other race students. Socioeconomic status (SES) was determined by qualification for free or reduced lunch. The looping team had 51% of students identified as low SES while 63% of the nonlooping teams' students were classified as low SES.

This school district has a history of allowing parent choice with regard to teacher selection. Approximately, 50% of the parents participate in this process at the middle school (Principal, personal communication, January 8, 2009). During the time of this study, parents were given the opportunity to select their child's teaching team with the knowledge that one particular team would loop. However, parents were as likely to select looping or nonlooping teams as the overall deciding factor was the teacher, not organization. Over the years, the school district has refined a process to disseminate students among teams based on parent choice, gender, ethnicity, and academic ability. Looping had been used previously in this district in lower elementary grades. This study coincides with the initial implementation of looping at the middle school. For the purposes of this study, students who were not promoted to a subsequent grade and students in special education who did not participate in statewide testing were excluded. MCT data for these two specific groups of students were not available, as such; they could not be included in the analysis.

Five mathematics teachers participated in the study. The looping teacher had a Bachelors of Science in Secondary Mathematics Education and a Masters of Arts in Elementary Education. This teacher had 10 years of teaching experience. The remaining four teachers taught either seventh or eight grade mathematics. One seventh-grade teacher had a Bachelors of Science in Mathematics and subsequently acquired teaching certification through an Alternate Route program; the other teacher had Bachelors of Science in Elementary Education. These teachers had 9 and 2 years of experience, respectively. Both eighth grade teachers had secondary certification and had earned Masters of Science in Curriculum and Instruction. The eighth grade teachers had more than 10 years of experience each.

Instrumentation

The MCT is a state designed criterion-referenced assessment that is given to all students enrolled in the public schools of Mississippi in grades 3 through 8 (Mississippi Department of Education, 2005). The Mississippi State Assessment Office has developed the series of assessments to measure student mastery of the state created curriculum. State-level curriculum specialists and professors of education, with the assistance of local teachers, reevaluate the curriculum framework every five years. The curriculum and assessments were created to meet the requirements set forth in the Elementary and Secondary Education Act (PL 89-10, 20 U. S. C.§ 6301 et seq.). Subsequently, state officials increased the number of grades assessed in response to specific mandates in *No Child Left Behind*.

The MCT test has three parts: reading, language, and mathematics. The MCT is a nontimed test given to the students on three consecutive days during the first week of May each year. The assessment used in this study is the mathematics component of the MCT, which is given on the third day of the testing cycle.

Procedure

The researchers chose a causal-comparative design to examine the effect of looping on student achievement in mathematics. The nature of this study drove the decision to choose this *ex post facto* approach. Participants were chosen from all students enrolled in the middle school from 2003–2005. Criteria for participation were the student was (1) enrolled in the middle school from 2003–2005; (2) was not retained for any of the years of the study; and (3) participated in the MCT each year. Researchers formed two groups of participants. Group one, the treatment group, consisted of students who looped in grades 7 and 8. Group two, the comparison group, consisted of students who did not loop. Descriptive statistics and MCT scores were collected on all participants. Finally, the effect size was calculated using Cohen's *d*.

Results

Data were collected to determine if middle school students taught by a looping team of teachers achieved higher mathematics scores on the MCT than middle school students who were not taught on a looping team. Mathematics MCT scores from 2003–2005 comprise this data set. Prior to performing the analysis to determine the effect of looping, it was necessary to ensure that the scores of the looping team and nonlooping teams were not statistically different in their initial placement in the seventh grade. Sixth grade MCT scores were used as baseline data. A one-way analysis of variance (ANOVA) was used to determine whether student scores were equitable. Next, a second ANOVA was completed to compare test scores and student growth for the teams. In addition, ANOVAs were performed to determine the statistical significance of gender, socioeconomic status, or ethnicity on scores and student growth (Gravetter & Wallnau, 2000; Howell, 2002). All data were evaluated using alpha level of significance set at .05. Finally, Cohen's d was determined to be $d = .40$ for this study, a small to moderate effect size.

Although the school district has a process for equally distributing students by gender, socioeconomic status, and ethnicity, the research team used sixth grade MCT mathematics scores to compare looping and nonlooping teams to verify team equality. Sixth grade MCT math scores averaged 586.87 for students on the looping team and 570.61 for students on nonlooping teams (see Table 1). The sixth grade math scores were not statistically significantly different between the two groups ($p = .054$) indicating that the looping and nonlooping groups were relatively evenly distributed.

Team Results

The average seventh grade MCT mathematics scores on the looping team were 612.42 compared with 586.39 for the nonlooping teams. The difference in

Table 1
Mississippi Curriculum Test Mathematics Results, Looping versus Nonlooping

	Sample *N*	Mean Score *M*			Change in Mean Score Δ		
Treatments		6th	7th	8th	6th–7th	7th–8th	6th–8th
Looping	69	586.870	612.420	621.970	25.550	9.550	35.100
Nonlooping	137	570.610	586.390	595.480	15.780	9.090	24.870
F-test (*p*)		.054	.002	.002	.041	.921	S.043

Table 2
Aggregated Mississippi Curriculum Test Mathematics Results: All Participants, Schools, and State

	Mean Score *M*			Change in Mean Score Δ		
Treatments	6th	7th	8th	6th–7th	7th–8th	6th–8th
School*	567.5	584.8	589.6	17.3	4.8	22.1
State	556.9	571.5	588.5	14.6	17.0	31.6

* Includes all students enrolled in the middle school.

mathematics scores between the teams in seventh grade was significant ($p = .002$). In the eighth grade, looping students averaged 621.97 compared with 595.48 for nonlooping students ($p = .002$). Thus, in seventh and eight grades, students in the looping classes scored higher on the mathematics portion of the MCT (see Table 1). Table 2 shows MCT scores as provided by the state for all tested students during the study. A comparison of looping team data to overall school data documents that the growth of looping participants from seventh to eighth grade was almost twice that of the schools' growth. Therefore, these results indicate that students in looping mathematics courses outperformed their counterparts in the nonlooping mathematics courses. Over the 2 years of the study, the looping students increased their mathematics MCT scores by an average of 35.10 points, whereas the nonlooping students increased their scores by an average of 24.87 points. This is statistically significant ($p = .043$). Also, a statistically significant difference in growth from sixth to seventh grade ($p = .041$), but not from seventh to eighth grade ($p = .921$) was found.

Gender

Females. There was no significant difference in the sixth grade MCT mathematics scores between the females who looped and those who did not loop, $p = .508$ (see Table 3). This indicates that the ability of the female students were evenly distributed across the looping ($M = 580.14$) and nonlooping teams ($M = 573.42$). The looping female students increased their MCT mathematics scores by 36.71 points, in comparison to 21.76 points for the nonlooping female students. The MCT seventh grade scores for looping female students increased 24.17 points compared to the 10.74 point increase for nonlooping females from sixth to eighth grade. The females who looped showed an increase of 12.54 points in eighth grade while the nonlooping females improved 11.01 points. This difference was not statistically significant ($p = .797$). However, overall growth from sixth to eighth grade for females was statistically significant ($F[1,115] = 5.584$, $p = .02$). During the seventh grade year, the female students on the looping and nonlooping teams averaged 604.31 and 584.16, respectively. The difference in scores between treatment and comparison was significant ($p = .044$). During the eighth grade year, female students who looped averaged 616.86 compared with 595.17 for nonlooping female students. There was no significant difference in scores between eighth grade looping and nonlooping females ($p = .051$).

Males. There was no significant difference in the sixth grade MCT mathematics scores between the males who looped and those who did not loop, $p = .56$ (see Table 3). This indicates that the ability of the male

Table 3
Mississippi Curriculum Test Mathematics Results by Gender

	Sample *N*	Mean Score *M*			Change in Mean Score Δ		
Treatments		6th	7th	8th	6th–7th	7th–8th	6th–8th
Female							
Looping	35	580.14	604.31	616.86	24.17	12.54	36.71
Nonlooping	82	573.42	584.16	595.17	10.74	11.01	21.76
F-test (*p*)		.51	.04	.05	.04	.80	.02
Male							
Looping	34	593.79	620.77	627.24	26.97	6.47	33.44
Nonlooping	55	566.42	589.71	595.93	23.29	6.22	29.51
F-test (*p*)		.06	.03	.02	.63	.61	.97

students were evenly distributed across the looping ($M = 593.79$) and nonlooping teams ($M = 566.42$). The seventh grade looping males mean scores were 620.77 in comparison to the nonlooping males means score of 589.71. The scores were found to be statistically significantly different ($F[1,87] = 3.766$, $p = .025$). In eighth grade, the looping males scored 627.24 in comparison to 595.93 for the nonlooping males. This difference was also statistically significant ($F[1,87] = 5.506$, $p = .021$).

There were statistically significant differences in mean scores between looping and nonlooping males. In seventh grade, the looping males had a mean score of 620.77 and nonlooping males had a mean score of 589.71, which was statistically significantly different ($p = .025$). Similarly, in eighth grade, the mean score for looping males was 627.24 and 595.93 for nonlooping males. This difference in mean scores was also statistically significantly different ($p = .021$). However, there were no statistically significant differences in growth over the two-year period as measured by change in MCT scores ($F[1,87] = .235$, $p = .629$). The looping males mean scores increased by 33.44 points, while the nonlooping males showed an increase of 29.51 points between sixth and eighth grades. Results are similar for differences in growth from the sixth to seventh and seventh to eighth grade years (see Table 3).

Ethnicity

African American and Caucasian looping students made greater gains over the two-year looping period than did their nonlooping peers. The African American students on the looping team gained 32.03 points in comparison to 23.96 points for the nonlooping African American students during the seventh and eighth grades. Total growth for looping Caucasian students was 36.74 points in comparison to 26.89 points for the nonlooping Caucasian students. Data analyzed for comparison between African American and Caucasian students were not statistically significant. During the study period, four students were classified as "other." Given this small N, they were removed from the sample (see Table 4).

Socioeconomic Status

Low socioeconomic (SES) students in both the looping and nonlooping groups scored significantly lower than the other students in their looping or nonlooping group, respectively. Low SES looping students began the seventh grade with their mean score 14.82 points below the mean scores of their low SES nonlooping peers. Growth occurred each year for low SES looping and nonlooping students. The amount of growth was not statically significant between low SES looping and nonlooping or within the looping group between the low SES group and other students. However, the looping students narrowed the gap between the mean scores of the low SES looping students and the low SES nonlooping students to 3.66 points (see Table 5).

Discussion

The purpose of this study was to examine the effect looping had on mathematics' achievement for seventh

Table 4
Mississippi Curriculum Test Mathematics Results by Ethnicity

	Sample *N*	Mean Score *M*			Change in Mean Score Δ		
Treatments		6th	7th	8th	6th–7th	7th–8th	6th–8th
African American							
Looping	36	555.11	573.00	587.14	17.89	14.14	32.03
Nonlooping	88	559.75	575.56	583.70	15.81	8.15	23.96
F-test (*p*)		.65	.79	.73	.73	.22	.17
Caucasian							
Looping	31	620.03	653.58	656.77	35.55	3.19	36.74
Nonlooping	47	589.98	573.00	616.87	15.19	11.70	26.89
F-test (*p*)		.02	.00	.00	.03	.37	.30

Table 5
Mississippi Curriculum Test Mathematics Results by Free/Reduced Lunch Status

	Sample *N*	Mean Score *M*			Change in Mean Score Δ		
Treatments		6th	7th	8th	6th–7th	7th–8th	6th–8th
Receiving Free/ Reduced Lunch							
Looping	35	545.40	567.89	579.11	22.49	11.23	33.71
Nonlooping	86	560.22	573.02	582.77	12.80	9.74	22.55
F-test (*p*)		.14	.57	.70	.12	.75	.07
Not Receiving Free/Reduced Lunch							
Looping	34	629.56	658.27	666.09	28.71	7.82	36.74
Nonlooping	51	588.12	608.92	616.90	20.80	7.98	26.89
F-test (*p*)		.00	.00	.00	.31	1.00	.38

and eighth grade students at one middle school. Given that the research base, both quantitative and qualitative, to support looping at the middle level is minimal, the researchers' secondary purpose was to contribute to this literature. Therefore, the progress of students in both looping and nonlooping teams was tracked using the mathematics portion of the state-wide criterion referenced test, the Mississippi Curriculum Test or MCT, during their seventh and eighth grade years. Scores on the MCT were compared generally and then across gender, ethnicity and socioeconomic status. The overall mathematics MCT mean score for each group was used as one data point. A second measure, growth in academic achievement as indicated by change in mathematics MCT scores, provides another method for analyzing the results of this study. The overall scores for both seventh and eighth graders on the looping team were statistically significantly higher than for students who did not loop. George and Shewey's (1997) and McLaughlin and Doda's (1997) work found that the looping environment fosters the development of long-term teacher–student relationships that leads to increased academic achievement. The results from this study support this finding with the overall scores of the looping students being significantly higher in both academic years. Despite the relatively small effect size, this study indicates there are significant interactions between mathematics and looping that need to be considered.

Comparisons among gender groups yielded interesting findings. The school intentionally distributed females and males evenly across the looping and nonlooping teams. Data indicated the performance levels on the MCT were also evenly distributed by gender. Despite the even distribution, females on the looping team scored statistically significantly higher on the MCT than their nonlooping counterparts in seventh grade. Further, females on the looping team showed statistically significant growth in their mathematics MCT scores from sixth to seventh and from sixth to eighth grade. Data for males also showed statistically significant differences in overall mathematics MCT scores in seventh and eighth grades. However, the growth for males was not significant. In other words, while looping males scored higher on the MCT, the growth from year to year was not significantly greater than their nonlooping counterparts. These findings may suggest that adolescent females tend to flourish when affect, a key component of looping, is found in their learning environment. Success in an academic area is a critical component in remaining interested in that particular area. As previously noted, students' intellectual support fostered in a nurturing learning environment are essential motivational sources in mathematics (National Mathematics Advisory Panel, 2008; Shores & Shannon, 2007). Our initial findings suggest that the looping environment supported females' mathematical growth and increased their mathematical competence; as a result, females might sustain interest in mathematics.

Over 65% of the school qualifies for free and/or reduced lunch. Because so many students are classified as low SES, the school did not use this criterion in placing students on looping or nonlooping teams. However, the researchers were interested in tracking the performance of this group. Data indicated that initially low SES students on the nonlooping teams had higher overall scores on the MCT (14.82 points higher). The low SES students on the nonlooping teams maintained the higher MCT scores over research years. Nonetheless, the low SES students on the looping team narrowed the gap on their MCT scores by 11.16 points. Researchers have found that the higher the level of poverty in a school, the lower the academic achievement is for all groups of students (Brown et al., 2003; Mertens & Flowers, 2003). The findings of this study support the belief that looping may create an environment that diminishes the influences of SES on academic achievement.

Perhaps the most striking finding from this research study is that all looping students showed greater gains in the first year of the looping cycle—not in the second. In seventh and eighth grades years, looping students had a larger increase in their academic achievement. The growth exhibited by looping students between sixth and seventh and between sixth and eighth grades was statistically significantly higher than those of the students in nonlooping teams. However, the growth from seventh to eighth grade was not statistically significant between the two groups. This finding may indicate that the initial investment made by teachers, students, and parents into the educational relationship is more profound in the first year of looping when everyone involved understands the long-term commitment. Consequently, teachers and students might exert more effort to develop a true community of learners which results in maximizing student performance in year one while maintaining academic achievement in year two. Further, this finding may suggest that students regain their academic interest and commitment to even challenging content areas, like mathematics, when they realize that their teachers have a vested interest in their academic well-being.

The work of George and Lounsbury (2000), George and Shewey (1997), and Jackson and Davis (2000) is built around the notion that middle school curriculum should be "relevant, challenging, exploratory, and integrated" (National Middle School Association, 2003). Mathematics educators share a similar vision for mathematics' classrooms. Each of the principles in Principles and Standards in School Mathematics (National Council of Teachers of Mathematics, 2000) can be connected back to the middle school tenets. Therefore, it is not surprising that the mathematical achievement of looping students was better than the achievement of their counterparts in traditional middle school classrooms. This study further supports the idea that motivation and intellectual support is a key component of success for both females and low SES groups of students (Kober, 2001; Shores & Shannon, 2007). Finally, STEM research suggests females selfselect out of STEM career fields in the middle grades, even though their mathematics performance is on-level with males. With the added intellectual and emotional support provided in a looping environment, females may find the desire to continue in traditionally male dominated fields.

Limitations

While the findings of this research study are intriguing, certain limitations to the study exist. First, this study was conducted in one southern school with a relatively small sample size. Future research would need to include multiple schools in both urban and rural settings. This need for further research is also indicated by the relatively small effect size. While the effect size is small, this may indicate there is importance in the interaction between looping and performance (Abelson, 1985). Also, investigating multiple looping teams in the same school should be considered to expand this research. A second limitation is that this study had only one mathematics teacher who taught in the looping environment. The current study did not allow for comparisons among this teachers pervious classes and the researched class, as historical data was unavailable. Further quantitative research into teacher effect and looping is warranted for all subject areas, not only mathematics.

References

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133.

Alexander, W. M. (1995). The junior high school: A changing view. *Middle School Journal*, *26*(3), 20–24.

American Association of University Women. (2004). *Under the microscope: A decade of gender equity projects in the sciences*. Washington, DC: American Association of University Women Educational Foundation. Retrieved January 16, 2009, from http://www.aauw.org/learn/research/upload/underthemicroscope.pdf

Brown, K. M., Roney, K., & Anfara Jr., V. A. (2003). Organizational health directly influences student performance at the middle level. *Middle School Journal*, *34*(5), 5–15.

Carraher, D. W., & Schliemann, A. D. (2007). Early algebra. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 669–706). Charlotte, NC: Information Age.

Comer, J. P. (2001). Schools that develop children. *The American Prospect*, *12*(7), 30–35.

Crosby, P. (1998). Looping in the middle school: Why do it? *Teaching PreK-8*, *29*(3), 46–47.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.

Darke, K., Clewell, B., & Sevo, R. (2002). Meeting the challenge: The impact of the National Science Foundation's Program for women and girls. *Journal of Women and Minorities in Science and Engineering*, *8*, 285–303.

Dentith, A. (2008). Smart girls, hard-working girls but not yet self-assured: The limits of gender equity politics. *Canadian Journal of Education*, *31*(1), 145–166.

Elliot, I. (1998). When two years are better than one. *Teaching PreK-8*, *29*(3), 38–41.

Fennema, E., & Leder, R. (1998). *Mathematics and gender: Influences on teachers and students*. New York: Teachers College Press.

Forsten, C., Grant, J., Johnson, B., & Richardson, I. (1997). *Looping Q & A: 72 practical answers to your most pressing questions*. Peterborough, NH: Crystal Springs Books.

Forsten, C., Grant, J., & Richardson, I. (1999). *The looping evaluation handbook*. Peterborough, NH: Crystal Springs Books.

George, P., & Alexander, W. (2003). *The exemplary middle school* (3rd ed.). Belmont, CA: Thomson-Wadsworth.

George, P., & Lounsbury, J. (2000). *Making big schools feel small: Multiage grouping, looping, and schools-within-a-school*. Westerville, OH: National Middle School Association.

George, P., & Shewey, K. (1997). What does research say? Maintaining long-term teacher student relationships. *Schools in the Middle*, *7*(2), 18–21, 58–59.

Good, T. L., & Brophy J. E. (2007). *Looking in classrooms* (10th ed.). New York: Allyn & Bacon.

Gravetter, F., & Wallnau, L. (2000). *Statistics for the behavioral sciences* (5th ed.). Stamford, CT: Wadsworth Thomson Learning.

Hayes Jacobs, H. (2004). *Getting results with curriculum mapping*. Alexandria, VA: Association for Supervision and Curriculum Development.

Howell, D. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury Thomson Learning.

Hyde, J. S. (1990). Meta-analysis and the gender psychology of gender differences. *Journal of Women in Culture and Society*, *16*(1), 55–73.

Jackson, A., & Davis, G. (2000). *Turning points 2000: Educating adolescents in the 21st century*. New York: Teachers College Press.

Kasak, D. (2004). What of the future? In S. Thompson (Ed.), *Reforming middle level education: Considerations for policymakers* (pp. 231–250). Greenwich, CT: Information Age.

Kellough, R. D., & Kellough, N. G. (2003). *Teaching young adolescents* (4th ed.). Upper Saddle River, NJ: Merrill Prentice-Hall.

Kieran, C. (2007). Learning and teaching of algebra at the middle school through college levels. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 707–762). Charlotte, NC: Information Age.

Kober, N. (2001). *It takes more than testing: Closing the achievement gap. A report of the center on education policy*. Washington DC: Center on Education Policy.

Manning, M., & Bucher, K. (2001). *Teaching in the middle school*. Upper Saddle River, NJ: Merrill Prentice-Hall.

McAteer, M. (2001). In the loop. *Teacher Magazine*, *13*(1), 17–19.

McCullough, C. (2003, June). *If we build it, will they come? Attracting and retaining under-represented groups in engineering*. Paper presented at American Society for Engineering Education Annual Conference, Nashville, TN.

McLaughlin, H., & Doda, N. (1997). Teaching with time on your side: Developing long-term relationships in schools. In J. Irvin (Ed.), *What current research says to the middle level practitioner* (pp. 57–72). Columbus, OH: National Middle School Association.

Mertens, S. B., & Flowers, N. (2003). Middle school practices improve student achievement in high poverty schools. *Middle School Journal*, *34*(1), 33–43.

Mertens, S. B., Roney, K., Anfara, V. A., Jr., & Caskey, M. M. (2007). *National middle grades research project: Common planning time training manual*. Washington, DC: Middle Level Education Research, a Special Interest Group of the American Educational Research Association.

Mississippi Department of Education. (2005). *MCT 2*. Retrieved November 14, 2005, from http://www.mde.k12.ms.us/acad/osa/mct2/

National Center for Education Statistics. (2003). *Trends in International Math and Science Study*. Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (2008). *Trends in International Math and Science Study*. Washington, DC: U.S. Department of Education.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards of school mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel. (2008). Foundations *for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.

National Middle School Association. (2003). *This we believe-successful schools for young adolescents; A position paper of the National Middle School Association*. Westerville, OH: National Middle School Association.

Nichols, J. D., & Nichols, G. W. (2002). The impact of looping classroom environments on parental attitudes. *Preventing School Failure*, *47*(1), 18–25.

Noddings, N. (1993). *Educating for intelligent belief and unbelief*. New York: Teachers College Press.

Payne, R. K. (2005). *A framework for understanding poverty* (4th ed.). Highlands, TX: aha! Process.

Powell, S. D. (2005). *Introduction to middle school*. Upper Saddle River, NJ: Pearson.

Shores, M. L., & Shannon, D. M. (2007). The effects of self-regulations, motivation, anxiety, and attributions on mathematics achievement for fifth and sixth grade students. *School Science and Mathematics*, *107*, 225–236.

Simel, D. (1998). Education for "*Bildung*": Teacher attitudes toward looping. *International Journal of Educational Reform*, *7*, 330–338.

Sleep, L., & Ball, D. L. (2007, January). *How does mathematical language figure in the work of teaching? How does this shape mathematical knowledge for teaching (MKT) and teacher education?* Presented at the Association of Mathematics Teacher Educators Annual Meeting, Irvine, CA.

Looping:

n A l n a o c

t e P e .

Florida International University

& n r a

Miami-Dade County Public Schools

Abstract: Looping is the practice in which a teacher instructs the same group of students for at least two school years, following them from one grade level to the next. Once a "loop" of two or more years is completed, the teacher may start a new loop teaching a new group of students. This evaluation study of the practice of looping in a large urban school system was intended to explore its effect on student instructional outcomes, attendance, and retention rates, as well as to assess principals' and teachers' reactions to looping. The results indicated that, with respect to academic achievement, the Looping Sample outperformed their counterparts in the Matching Sample. Looping had a positive effect on student attendance and students in the Looping Sample had a significantly greater chance of being promoted to the next grade level. Principals and teachers were in high agreement that looping had a positive effect on student learning in their schools.

Looping: An Empirical Evaluation

Looping is the practice in which a teacher instructs the same group

Address all correspondence to: Peter J. Cistone, Department of Educational Leadership and Policy Studies, Florida International University, Miami, Florida 33199. E-mail: cistonep@fiu.edu

International Journal of Educational Policy, Research, & Practice, Volume 5, Number 1, Spring 2004. ISSN 1528-3534; Copyright 2004 by Caddo Gap Press; All rights of reproduction reserved.

of students for at least two school years, following them from one grade level to the next. Once a “loop” of two or more years is completed, the teacher may start a new loop teaching a new group of students. The practice of looping has been described under various names, including teacher rotation, family-style learning, student-teacher progression, and multiyear instruction.

Looping has been employed in education for some time. Rudolf Steiner founded the Waldorf Schools in Germany in the early 1900s on the notion that students would benefit from a lasting relationship with a teacher. In Waldorf Schools, teachers remained with their students during grades one through eight. This practice continues today in the Waldorf Schools that have expanded to many countries around the world. Currently, in Germany, students and teachers generally stay together in grades one through four (Northeast and Islands Regional Educational Laboratory at Brown University, 1997). Looping is also practiced in other countries, including Israel, Sweden, and Japan. In these countries, looping is used by many schools in elementary grades. Modified versions of multi-year teacher-student relationships are in place in secondary grades as well (Grant et al. as cited in Little and Little, 2001). Preschools in Italy successfully use a three-year teacher-students assignment model (Palestis, 1994).

In the United States, Deborah Meier, a well-known New York City educator and author, began using looping in 1974. She reasoned that teachers and students needed to become well-acquainted with one another in order to achieve necessary levels of communication that would support learning. Meier considers looping important in providing teachers and students with an opportunity to get to know each other very well (Goldberg, 1991). Deborah Jacoby, another looping practitioner and supporter, describes the time saved on the assessment of skills, increased ability to utilize the children’s known strengths and talents in a style consistent over two years, and trusting relationships built with students and parents as some advantages of looping (Jacoby, 1994).

The literature on looping reports many benefits of this practice. Looping allows teachers to save time at the beginning of the second year of the loop by making unnecessary the usual transitional period typically spent on getting acquainted with new students as well as setting classroom rules, expectations, and standards. The time saved is virtually identical to gaining an extra month of teaching/learning time during the second year of the loop (Burke, 1996; Black, 2000).

Moreover, research indicates that looping gives children more time to build relationships essential to learning and aids in the development of social skills (Checkley, 1995), reduces anxiety experienced by students when they go from one grade level to the next (Grant & Johnson, 1995), and improves student confidence and parent-teacher relationships (Little & Dacus, 1999). Teachers in looping classes develop a closer relationship with students' parents (Rasmussen, 1998) and the practice of looping positively affects parents' attitudes toward the educational environment (Nichols & Nichols, 2002). There is evidence that looping may serve to improve overall elementary school climate (Black, 2000).

Concomitantly, the literature on looping indicates some potential disadvantages of looping. If the teacher is not familiar with the curriculum of the second year of the loop, the valuable instructional time may be lost. There can be a mismatch between teaching style and a child's learning style. Going forward with this mismatch for more than one school year is bad for both the teacher and the student. With looping, a student may have to be taught by an instructor who is not very strong in a particular subject area for more than one year (Vann, 1997). Others who have studied looping suggest that some of these drawbacks of looping may prove to be advantages. Chapman (1999) states that the problems concerning teacher/student mismatch or weak teachers should be addressed by a principal in any case—not just in looping situation. Looping may encourage principals to act more strongly to address these problems.

The findings concerning benefits of looping mostly reflect its social advantages for students. There appears to be a paucity of recent empirical studies targeting the academic effects of looping, especially its effects on student academic achievement. The present study aims to address the academic effects of looping.

Within a large urban school system in the state of Florida, 26 elementary schools used looping in the 1999-2000 school year. In these schools, looping was implemented in a variety of ways. In certain schools, only gifted students or students in the Advanced Academic Placement program participated in looping, while in others, students in regular classes took part in it. In some schools, only one or two classes participated in looping, whereas in others, all classes in particular grade levels took part in it. In addition, looping patterns were organized differently among schools. In certain schools, the looping occurred in first and second grades, and then in third and fourth grades, while in some other schools it was implemented in the second and third grades only.

This evaluation of the practice of looping was intended to explore its effect on student instructional outcomes, attendance, and retention rates, as well as to assess principals' and teachers' reactions to looping.

Method

This study intended to explore the academic effects of looping regarding general education students (as opposed to gifted or advanced placement students) who completed a two-year “loop.” All 26 elementary schools that used looping during the 1999-2000 school year were considered. Of the 26 schools, 11 were in the first year of the loop or had only gifted or Advanced Academic Placement program students participating in looping. Accordingly, these schools were excluded. Then, for the purpose of making necessary comparisons, two student samples were created. These two samples represented students participating in looping and matching peers not participating in it. Clearly, it was necessary to assure that students in the two samples were similar in terms of their demographic characteristics and academic achievement before the looping began—in the 1997-1998 school year.

Looping Sample. Since looping is a multiyear program, only students who were in this program for its entire duration could reap all its benefits. Thus, the Looping Sample included all those students from looping classes of selected schools who were taught by the same teacher during the 1998-1999 and 1999-2000 school years. This sample consisted of 612 students.

Matching Sample. The Matching Sample was created in two stages. First, students in the Looping Sample were matched to those students of non-looping schools in the school system who stayed in the same school during 1998-1999 and 1999-2000 school years and who matched the students in the Looping Sample in terms of gender, race/ethnicity, status on free/reduced lunch, primary exceptionality, and English for Speakers of Other Languages (ESOL) level. This procedure created a group of possible “matches” for each student in the Looping Sample. Then, for each student in the Looping Sample, the results in reading comprehension and mathematics applications on the 1998 Stanford Achievement Test, Eighth Edition, were used to choose one person who best matched the student in the Looping Sample in terms of academic achievement. (The closeness of the match was established by minimizing the sum of the squared deviations of mathematics and reading test scores from those of the student in the Looping Sample.) Most of the students in the Matching Sample (410) were matched to their counterparts in the Looping Sample using both the demographic and the achievement criteria above. The remaining 202 students did not participate in the 1998 Stanford Achievement Test due to their grade levels; therefore, they were matched to their peers in the Looping Sample on all of the demographic parameters listed above.

The closeness of the academic achievement match between students in the two samples above is evident from a comparison of the 1998 achievement results (prior to the beginning of a two-year loop). The mean scaled score in mathematics was 611.5 for the students in the Looping Sample, which was very close to 612.1, the mean scaled score for their counterparts in the Matching Sample. Similarly, the mean scaled score in reading was 606.6 for students in the Looping Sample and 606.7 for their peers in the Matching Sample, an almost identical figure. The demographic characteristics of students in both samples are shown in Table 1.

*Academic Achievement Comparisons***.** The norm-referenced component of the Florida Comprehensive Assessment Test (FCAT) administered in March 2000 was used to compare the academic achievement of students in the Looping and Matching Samples. (The state of Florida used a special edition of the Stanford Achievement Test, Ninth Edition, as the norm-referenced component of the FCAT.) Since the mean achievement results in reading and mathematics obtained before the beginning of the loop (in March of 1998) were virtually identical for students in the two samples, no statistical adjustment for prior achievement was necessary. Consequently, the paired-samples t-test was used to statistically compare the reading and mathematics achievement outcomes for students in the Looping and Matching Samples. This test requires the data be available for each student in a matched pair of students. Mathematics applications results were available for 581 matched pairs of students, and reading comprehension results were available for 577 matched student pairs. (The rest of the students either did not participate in the norm-referenced FCAT or their tests were invalidated.) Therefore, 581 and 577 paired achievement scores in mathematics and reading respectively were used for academic comparisons of students in the Looping and Matching Samples.

Attendance and Retention Comparisons. The end-of-year data for the 1999-2000 school year were used to compare the attendance and retention rates of students in the Looping and Matching Samples. These data were available for all 612 students in both samples. The paired-samples t-test was used to compare the differences in the average number of days absent

Table 1
Student Demographic Characteristics

	Race/Ethnicity				Gender	
	White Non-Hisp.	Black Non-Hisp.	Hisp.	Other	Female	Male
Grade 2 (n = 185)	9%	2%	87%	2%	52%	48%
Grade 3 (n = 28)	0%	68%	21%	11%	50%	50%
Grade 4 (n = 296)	9%	12%	77%	2%	55%	45%
Grade 5 (n = 103)	7%	50%	43%	0%	52%	48%
Total (n = 612)	8%	18%	72%	2%	54%	46%

between the 1998-1999 and 1999-2000 school years for students in the two samples. The log odds analysis was conducted to compare the numbers of students not promoted to the next grade level in the two groups.

Principals' Survey. Principals of the elementary schools that used looping in the 1999-2000 school year were surveyed. The Principal Questionnaire consisted of eight true-false questions designed to measure respondents' opinions about the benefits of looping and three open-ended questions asking principals to describe the criteria for selecting teachers to be involved in looping, and the advantages and shortcomings of looping as it was implemented in their schools. The principals of all 26 elementary schools in which looping took place during the 1999-2000 school year were asked to complete the survey. Eighteen of them returned completed questionnaires (69% return rate).

Teachers' Survey. All teachers in the school system who were involved in looping were surveyed. The Teacher Questionnaire consisted of two parts. The first part, containing 14 true-false items, was intended to measure respondents' reactions to looping; the second part, which consisted of two open-ended items, asked teachers to describe the advantages and shortcomings of looping as it was implemented in their schools. In all, 96 teachers were asked to participate in the survey; 69 of them returned completed questionnaires (72% return rate). However, only 58 questionnaires were used for the analysis (60% rate), because the remaining 11 teachers were in their first year of the loop and did not participate in the looping in the past. It is generally believed that most of the benefits of looping are realized during the second year of the loop, which implies that these teachers were not in the position to answer most of the questions about the benefits of looping.

It should be noted that it was not possible to select teachers randomly for participation in looping, nor was it possible to assign teachers or students randomly to looping and non-looping classes. Consequently, the findings reported below should be understood accordingly.

Results

The comparison of academic achievement for students in the Looping Sample and their counterparts in the Matching Sample was based on the results of the norm-referenced part of the FCAT administered in March 2000. The outcomes are presented separately for the reading comprehension and mathematics applications sections of the test.

Reading Achievement Results. Since the number of students in each grade level was the same for both samples, it was possible to compare the scaled scores for the two samples across all grade levels. A paired-sample t-test was performed to determine whether the students in the Looping Sample, as a group, scored significantly higher on the reading comprehension section of the FCAT than did students in the Matching Sample. The results indicated that the mean scaled score for students in the Looping group ($\underline{M}$ = 634, $\underline{SD}$ = 42) was significantly greater than that for students in the Matching group ($\underline{M}$ = 628, $\underline{SD}$ = 44), $\underline{t}(576) = 3.78$, $\underline{p} < .001$. The 95% confidence interval of the difference scores was (2.93, 9.29). The standardized effect size index $\underline{d}$ was .16, a value generally considered small. The magnitude of the effect size index indicates that an average student in the Looping Sample outperformed about 56% of students in the Matching Sample on the reading comprehension part of the FCAT. In terms of the raw scores, an average student in the Looping group answered two to three more multiple-choice questions correctly than did an average student in the Matching group. (The maximum raw score was between 40 and 54 points depending on the grade level.)

Furthermore, the students in the Looping Sample consistently outperformed the students in the Matching Sample on the reading comprehension section of the FCAT across the different grade levels that the samples comprised. The mean reading scaled scores of students in the Looping Sample were higher than those of students in the Matching Sample for all grade levels. The reading achievement results, expressed as percentile ranks corresponding to the mean scaled scores for students in both samples, are presented in Table 2.

It can be seen that students in the Looping Sample on average have substantially higher percentile scores than do their counterparts in the Matching Sample. The difference in percentile ranks that correspond to mean scaled scores for students in the two samples varies from four to

Table 2
Reading Achievement Results on the FCAT (Norm-Referenced Test) by Grade Level

	Percentile Corresponding to the Mean Scaled Score		Difference in Percentile Scores
	Looping Sample	Matching Sample	Looping/ Matching
Grade 2 ($\underline{n}$ = 159)	68	63	+5
Grade 3 ($\underline{n}$ = 27)	61	53	+8
Grade 4 ($\underline{n}$ = 293)	63	58	+5
Grade 5 ($\underline{n}$ = 98)	38	34	+4

Note: Some of the percentiles are interpolated.

eight percentile points, a sizable amount. This effect is consistent across all grade levels included in the samples.

Mathematics Achievement Results. The mathematics applications section of the norm-referenced part of the FCAT was used to make academic achievement comparisons for students in the Looping and Matching Samples. A paired-sample t-test was performed to determine whether the students in the Looping Sample, as a group, scored significantly higher on the mathematics applications section of the FCAT than did students in the Matching Sample. The results indicated that the mean scaled score for students in the Looping group ($\underline{M}$ = 628, $\underline{SD}$ = 39) was significantly greater than that for students in the Matching group ($\underline{M}$ = 620, $\underline{SD}$ = 42), $\underline{t}$ (579) = 4.95, $\underline{p} < .001$. The 95% confidence interval of the difference scores was (4.68, 10.83). The standardized effect size index $\underline{d}$ was .21, a value generally considered small. The magnitude of the effect size index indicates that an average student in the Looping Sample outperformed about 58% of students in the Matching Sample on the mathematics applications part of the FCAT. An average student in the Looping group answered two to three more multiple-choice questions correctly than did an average student in the Matching group. (The maximum raw score was between 46 and 48 points depending on the grade level.)

Moreover, students in the Looping Sample outperformed their peers in the Matching Sample on the mathematics application section of the FCAT across all grade levels represented in both samples. The mean mathematics scaled scores of students in the Looping Sample were higher than those of students in the Matching Sample across all grade levels. The results of these comparisons, expressed in terms of percentile ranks corresponding to the mean scale scores, are presented in Table 3.

Table 3
Mathematics Achievement Results on the FCAT (Norm-Referenced Test) by Grade Level

	Percentile Corresponding to the Mean Scaled Score		Difference in Percentile Scores
	Looping Sample	Matching Sample	Looping / Matching
Grade 2 ($\underline{n}$ = 163)	71	64	+7
Grade 3 ($\underline{n}$ = 27)	69	61	+8
Grade 4 ($\underline{n}$ = 292)	66	57	+9
Grade 5 ($\underline{n}$ = 98)	59	53	+6

Note: Some of the percentiles are interpolated.

It can be seen that students in the Looping Sample have substantially higher percentile ranks on the mathematics applications section of the FCAT than do their peers in the Matching Sample. The difference in performance expressed by percentile ranks corresponding to the mean scaled scores for students in the two samples varies from six to nine percentile points—a considerable amount. This effect is consistent for all grade levels that the samples comprise.

Attendance Comparisons. The attendance of students in the Looping and Matching Samples was compared. As mentioned earlier, students in the Looping and Matching Samples were equated on several demographic characteristics and matched on academic performance measured prior to the beginning of the loop. However, students in the two samples were not matched on the absenteeism figures. As shown in Table 4, the average numbers of days absent during the 1998-1999 school year (the first year of the loop) were different: approximately eight for students in the Looping Sample and seven for their matching counterparts. Since students in the two samples had different attendance levels during the first year of the loop, it was necessary to examine the differences (increases or decreases) in the average numbers of days absent between the second and first years of the loop for students in the Looping and Matching Samples. A paired samples t-test was performed to determine whether a decrease in the average number of days absent between the 1998-1999 and 1999-2000 school years for students in the Looping Sample was greater than that for students in the Matching Sample. The results showed that the decrease in the mean number of days absent for students in the Looping group ($\underline{M}$ = .78, $\underline{SD}$ = 5.14) was significantly greater than that for students in the Matching group ($\underline{M}$ = –.18, $\underline{SD}$ = 5.71), $\underline{t}$ (611) = 3.08, $\underline{p}$= .001. The 95 % confidence interval of the difference in the decrease of the number of days absent was (.35, 1.57). The standardized effect size index $\underline{d}$ was .12, a small value.

Students in almost all grade levels represented in the Looping Sample exhibited improved attendance. The average number of days absent decreased by approximately one or two days for students in the second, third, and fifth grades between the two academic years and remained at virtually the same level for the fourth graders. The actual absenteeism figures during the 1998-1999 and 1999-2000 school years for the two student groups are shown in Table 4.

The evidence collected indicate that students in the Looping Sample improved their attendance from one academic year to the next, while the attendance levels of students in the Matching Sample decreased during the same period. This fact suggests that looping had a positive effect on student attendance.

Table 4
Average Number of Days Absent

	Looping Sample			Matching Sample		
	1998-1999 School Year	1999-2000 School Year	Increase/Decrease	1998-1999 School Year	1999-2000 School Year	Increase/Decrease
Grade 2 (n = 185)	9.4	8.2	–1.2	9.1	8.3	–.8
GRADE 3 (n = 28)	7.4	5.0	–2.4	5.5	5.4	–.1
Grade 4 (n = 296)	7.2	7.0	–.2	6.5	6.9	+.4
Grade 5 (n = 103)	8.0	6.7	–1.3	6.5	7.8	+1.3
Total (n = 612)	8.0	7.2	–.8	7.2	7.4	+.2

Note: The grade levels shown are for the 1999-2000 school year.

Student Retention Comparisons. The retention figures for students in the Looping Sample in the 1999-2000 school year (the second year of the loop) and that of their counterparts in the Matching Sample were compared. The overall retention figures show that only two students in the Looping Sample were retained as compared to seven students in the Matching Sample (see Table 5). A log odds analysis was conducted to determine whether the number of students in the Looping group who were held back at the end of 1999-2000 school year was smaller than the corresponding figure for students in the Matching group. The two variables were group with two levels (Looping or Matching) and student status with two levels (promoted to the next grade level or held back). The results showed that the odds ratio was 3.53, indicating that a student in the Looping group was 3.53 times more likely to be promoted to the next grade level than a student in the Matching group. When the log odds test of significance was performed, it was found that at the common .05 level of significance there was not sufficient evidence to say that the students in the Looping Sample had a significantly different chance of being promoted to the next grade level than students in the Matching Sample

Table 5
Retention Results of the 1999-2000 School Year by Grade Level

	Promoted to the Next Grade	Held Back	Odds of Promotion	Odds Ratio
Looping Sample	610	2	305	3.53
Matching Sample	605	7	86.43	

(p = .06). Of course, the same result would be considered significant at a less stringent .1 level.

Teacher and School Principal Surveys. Eighteen principals and 60 teachers of looping classes were surveyed. The results show that most participants believed that looping had a positive effect on students' learning in their schools.

Proponents of looping usually state that teachers in looping classes gain some learning time at the beginning of the second year of the loop, because they do not need to learn their students' names, personalities, and learning styles. In our survey, almost all principals (94%) and most teachers (91%) agreed with this statement and indicated that looping increased the time available to teachers at the beginning of the second year of the loop (see Table 6). Another advantage of looping asserted by its supporters is that it increases the time available to slower students to learn the basic skills. Most of the principals (89%) and the majority of the teachers surveyed (71%) agreed with his assertion. In addition, most principals (89%) and the majority of teachers surveyed (72%) stated that looping enhanced the working relationship between teachers and students. Finally, most principals (94%) and nearly all teachers (95%) indicated that overall, looping increased the effectiveness of classroom instruction.

There was one area, however, where the teachers' opinions differed

Table 6
Opinions about Looping

	Percent in Agreement	
	Principals (n = 18)	Teachers (n = 60)
Looping increases the instructional time available to teachers at the beginning of the second year of the loop.	94%	91%
With looping, slower students have more time to learn the basic skills.	89%	71%
Looping enhances the quality of the working relationships between teachers and students.	89%	72%
Looping increases parental involvement in education.	72%	46%
Overall, looping enhances the effectiveness of classroom instruction.	94%	95%

from principals'. More than half of the principals (72%) stated that looping raised parental involvement in education, but only 46% of the teachers agreed with this statement.

In addition to the questions that were posed to both the principals and teachers and presented in Table 6, there were some questions that only principals or teachers were asked. The replies to the questions directed to principals revealed that most principals (89%) believed that looping decreases the number of disciplinary problems in schools. A small number of principals (11%) thought that only experienced teachers should teach looping classes. Only 11% of principals surveyed stated that they often had to deal with student-teacher or parent-teacher personality conflicts.

When teachers were asked a similar question, 37% of them indicated that looping increases the chance of student-teacher personality conflict. The majority of teachers (75%) were concerned that, with looping, teachers sometimes have to deal with an unreasonable parent for a long time, but at the same time, 83% of teachers stated that parents of students in looping classes usually have good working relationships with teachers. All teachers stated that looping enables teachers to accumulate detailed knowledge about their students, and most teachers (93%) indicated that looping helps them to individualize instruction. The majority of teachers (88%) believed that looping increases time available to slower students to learn the basic skills. Nearly all teachers (98%) indicated that students in looping classes feel less apprehensive at the beginning of the second year of the loop. Teachers also strongly believed that they should be allowed to choose whether to participate in looping. All teachers surveyed expressed this opinion. However, this conviction does not imply that the majority of the teachers had reservations about participating in looping. In fact, just the opposite was the case: the majority of the teachers surveyed (81%) stated that, given a choice, they would like to teach a looping class again.

In addition to responding to the true/false questions, most principals (89%) and the majority of teachers (75%) provided comments about looping as it was implemented in their schools. Most of the comments described advantages of looping. Remarkably, all 16 principals who commented on looping stated that they did not see any drawbacks to this practice. A number of teachers who commented on the program expressed the same opinion. Very few comments from teachers addressed the disadvantages of looping. Only four teachers voiced concerns about possible student-teacher personality conflict, and three teachers pointed out parent-teacher misunderstanding as a detriment to looping. By contrast, 25 teachers indicated that looping allowed them to gain in-depth knowledge about their students' academic strengths and weaknesses, personalities, and learning styles. This knowledge, in turn, allowed teachers to start instructional activities immediately at the beginning of the second year of the loop; no time was spent getting acquainted with students. One teacher wrote, "I really like the head start looping allows one to have during the course of the new school year. Personalities are known and personal relationships have been established. As a result, time on task is increased and behavior problems are minimized."

Several principals surveyed commented on the way teachers are selected to work with looping classes. The majority of the principals who provided these comments indicated that they selected teachers based on their requests. In two elementary schools all classes in grades one and two, and then three and four participated in looping.

The principals' and teachers' survey results indicate that the majority of participants in both groups had positive attitudes toward looping. The majority of respondents in both groups stated that looping provided more time to slower students to learn basic skills. Most respondents indicated that looping allowed teachers to gain learning time at the beginning of the second year of the loop, and nearly all principals and teachers surveyed stated that looping enhanced the effectiveness of the classroom instruction. In addition, although all teachers believed that they should be given a choice on whether to participate in looping, most teachers surveyed indicated that, given a choice, they would like to participate in looping again.

Conclusions

Findings based on analyses of student academic performance, retention and absenteeism figures, and teacher and principal surveys indicate that looping has a beneficial educational effect on students, and that it is viewed positively by school personnel.

The results of the analyses of student academic achievement demonstrate that students in the Looping Sample, as a group, exhibited significantly higher academic performance on the reading comprehension and mathematics applications sections of the FCAT than did students in the Matching Sample. Furthermore, students in the Looping Sample substantially outperformed their matched counterparts in both areas across all grade levels included in the samples. These facts suggest that participation in looping increased student academic achievement.

The result of the analysis of student's absenteeism figures shows that students in the Looping Sample, as a group, improved their attendance between the first and second years of the loop. The average attendance of Matching Sample students declined during the same period. This finding suggests that participation in looping improved student attendance.

The results of the student retention figures demonstrated that the number of students in the Looping Sample retained after the 1999-2000 school year was significantly lower than the corresponding figure for the Matching Sample. This suggests that participation in looping reduced student retention.

Most principals and teachers surveyed had positive opinions about looping. The majority of respondents agreed that looping enhanced a working relationship between teachers and their students. Furthermore, most teachers and principals surveyed stated that looping provided more time for slower students to learn basic skills. Moreover, almost all respondents indicated that, with looping, teachers can gain learning time at the beginning of the second year of the loop, and nearly all respondents stated that looping had a positive impact on learning in their schools. Most teachers surveyed were enthusiastic about looping. Although all teachers surveyed believed that teachers should be allowed to decide whether to participate in looping, most of them stated that they would like to participate in looping again. Finally, principals' and teachers' replies to the questionnaires indicated that, in their opinion, the benefits of looping greatly outweighed its drawbacks.

These findings suggest that looping can become a feasible school restructuring choice providing valuable educational benefits without significantly increasing operational costs.

References

Black, S. (2000). Together again. *The American school board journal, 187(6)*, 40-43.

Burke, D. L. (1996). Multi-year teacher/student relationships are a long-overdue arrangement. *Phi Delta Kappan, 77(5)*, 360-361.

Chapman, J. (1999). A looping journey. *Young children, 54(3),* 80-83.

Checkley, K. (1995). Multiyear education: Reaping the benefits of "looping." *ASCD Education Update, 37(8)*, 1-6.

Goldberg, M. F. (1991). Portrait of Deborah Meier. *Educational Leadership, 48*, 26-28.

Grant, J. & Johnson, B. (1995). Looping, the two-grade cycle: A good starting place. In *A Common Sense Guide to Multiage Practices, Primary Level* (pp. 33-36). Columbus, OH: Teacher's Publishing Group.

Jacoby, D. (1994). Twice the learning and twice the love. *Teaching Pre K-8, 24(6)*, 58-59.

Little, T. S. & Dacus, N. (1999). Looping: moving up with the class. *Educational Leadership, 57(1)*, 42-45.

Little, T. S., & Little, L. P. (2001). Looping: creating elementary school communities. *Phi Delta Kappa Fastbacks, 478,* 7-39

Northeast and Islands Regional Educational Laboratory at Brown University. (1997, November). *Looping: Supporting Student Learning Through Long-Term Relationships.* Providence, RI.

Nichols, J. D. & Nichols, G. W. (2002). The impact of looping classroom environments on parental attitudes. *Preventing School Failure, 47(1),* 18-22.

Palestis, E. (1994). Lessons from Reggio Emilia. *Principal, 73(5),* 16-19.

Rasmussen, K. (1998). Looping—discovering the benefits of multiyear teaching. *Association for Supervision and Curriculum Development, 40(2),* 2-4.

Vann, A. S. (1997) Looping: looking beyond the hype. *NAESP—Principal Magazine,* from http://www.naesp.org/comm/p0597c.htm