New York City Department of Education Literacy Intervention Pilot Project
In collaboration with Reading Reform Foundation and the Rose Institute for Learning and Literacy at Manhattanville College

Final Report on Student Progress 2015 - 2016

Prepared by Jolie Heath and Linnea C. Ehri
Graduate Center of the City University of New York

Introduction to the Project

In the spring of 2015, Reading Reform Foundation of New York and the Rose Institute for Learning and Literacy at Manhattanville College were asked by the New York City Department of Education to engage in a collaboration to develop and implement a literacy intervention pilot project during the 2015-16 school year in seven schools across the five boroughs of New York City. In this project, 20 classroom teachers and one AIS teacher in kindergarten through second grade received intensive professional development through course work and in-class mentoring in an integrated language-arts program, *The Writing Road to Reading,* based on the work of Romalda Spalding.

Before the school year 2015-16 began, the participating teachers took a two-week summer course on the Spalding approach provided by Reading Reform Foundation and the Rose Institute. This course provided the foundational strategies of the approach.

From September 2015 until May of 2016, project coaches met with each teacher on two separate days during the week for individual, private lesson planning and practice (40+ minutes) and in-class implementation of the planned lessons, with practice, modeling and coaching (40+ minutes) for an approximate total of two hours and forty minutes each week.

Pretest Measures

In the fall, before classroom instruction began, several measures were administered to students of the 20 teachers to assess beginning reading and spelling skills across the grades:

In kindergarten:

- The Sands Sound/Symbol test was administered to assess children's ability to recognize fifteen letters and their sounds and to write five words (*am, at, mat, dot, sit*).

- The Gates-MacGinitie subtest of letters and sounds (Level K, Form PR) containing thirty items further assessed alphabetic knowledge.

- The Morrison McCall Spelling Scale was first administered as a pretest in March, 2016, rather than at the beginning of the school year since many children entering kindergarten have limited handwriting and spelling skills. On this task, students spelled the first ten words from Morrison McCall List # 6.

In first grade:

- Children were administered the Gates-MacGinitie subtest of word decoding (Level 1, Form S) containing forty-three items.

- The Morrison McCall Spelling Scale was given and students were asked to spell the first thirty words from List # 1.

In second grade:

- Students were administered the Gates-MacGinitie subtest of word decoding (Level 2, Form S) containing forty-three items.

- Students were administered the Gates-MacGinitie subtest of reading comprehension (Level 2, Form S) containing thirty-nine items.

- The Morrison McCall Spelling Scale was administered. Students were asked to spell the first forty words from List # 1.

Posttest Measures

In May, after almost a full school year of classroom instruction, several measures were administered again to assess growth in beginning reading and spelling skills across the grades:

In kindergarten:

- The Sands Sound/Symbol test was given to assess recognition of fifteen letters and their sounds and the ability to write five words (*am, at, mat, dot, sit*).

- The Gates-MacGinitie subtest assessed knowledge of letters and sounds (Level K, Form PR) with 30 items.

- The Morrison McCall Spelling Scale required students to spell the first 10 words from List # 1.

In first grade:

- The Gates-MacGinitie subtest assessed word-decoding ability (Level 1, Form S) with 43 items.

- The Morrison-McCall Spelling Scale required students to spell the first 30 words from List # 1.

In second grade:

- The Gates-MacGinitie subtest assessed word-decoding ability (Level 2, Form T) with 43 items.

- The Gates-MacGinitie subtest of reading comprehension (Level 2, Form T) included 39 items.

- The Morrison-McCall Spelling Scale required students to spell the first 40 words from List # 1.

Purpose of this Report

This report presents evidence regarding the progress of students in kindergarten, first and second grades who received the collaboration's intervention between September 2015 and May 2016. It examines student achievement as the result of both professional development and ongoing coaching to improve teachers' effectiveness teaching beginning reading and spelling with the Spalding approach.

Study Method

- The participants were 20 teachers and their students ($N = 391$) across three grades: kindergarten ($N = 72$), first grade ($N = 201$) and second grade ($N = 118$). All teachers received training in the summer (45 hours) and in-class instructional support throughout the year (approximately 104 hours) provided by the Reading Reform Foundation and Rose Institute mentors ($N = 6$).

- Comparative data for all the kindergarten, first, and second grade students who completed all pretest and posttest measures are included in the analyses. The average performance of each classroom is reported at two test points to reveal how much students improved on average over the course of the school year from the instruction they received.

- Kindergarteners were assessed in September and May on two tests of letter-sound knowledge and word spelling. In addition, kindergarteners were further assessed with a second spelling measure as an interim test in March, 2016. This spelling measure was administered again as a posttest in May. The average performance of students in each classroom is reported at two test points to measure student gains in early literacy acquisition of letter-sound knowledge and spelling.

- First graders were assessed in September and May on two measures: word reading and spelling. The average performance of each classroom at the beginning and end of the study on these measures is reported.

- Second graders were assessed in September and May on three measures: word reading, reading comprehension, and spelling. The average performance of each classroom at the beginning and end of the school year on these measures is reported.

- Coaches reported on teacher and class progress throughout the school year in monthly logs. These surveys documented competencies and skills that teachers were expected to master during the school year as a result of participating in the RRF program: teacher attitude, mastery at pronouncing sounds accurately, introducing new sounds, dictation of words, teaching of handwriting and spelling rules. In addition, these logs recorded whether teachers were able to present lessons during each month, whether observations of lessons were made, and any areas of challenge or struggle that a teacher may have faced.

End-of-year Results

The average test scores for each classroom of students in both September and May are displayed in Tables 1 and 2 (kindergarten), 3 and 4 (first grade), and 5 and 6 (second grade) along with characteristics of schools and classes. Mean scores provide an estimate of reading and spelling performance of students in each class and are presented in two ways: 1) as mean raw scores (RS) indicating the average number of correct items that students in a given class scored on a test, and standard deviations in parentheses showing the average spread, that is, the average difference between the mean score and individual scores, and 2) as mean grade equivalent (GE) scores determined by standardized norms reported in test manuals to allow comparison to expected performance levels of students at that grade level. For some of the kindergarten tests, GEs were not provided by test manuals. Superscripted letters indicate explanatory information appearing in footnotes.

Tables 1 through 6 report the mean scores for kindergarten, first grade, and second grade class performance on all literacy measures. The mean scores of classes were greater on May posttests than on fall pretests or on the March interim test in kindergarten. To verify this, paired sample statistical t-tests were conducted. Results revealed statistically significant growth from pretest to posttest, indicating that all classes showed impressive gains on all measures. These gains are depicted in Figures 1 through 9, which display progress from September to May for each measure and each classroom teacher.

In kindergarten:

Tables 1 and 2 report mean scores for kindergarten classes on four measures, two separate measures of letter-sound knowledge, and two separate measures for spelling words. Overall, kindergarten students' mean recognition of letter sounds and mean spelling performance was greater in May than in September and March. These gains are depicted in Figures 1 through 4 for each measure.

In Table 1 mean scores are shown for letter-sound knowledge. In the fall, students were able to recognize more than half of the sounds of letters on this measure. By May, kindergarteners' mean performance on letter-sound recognition grew substantially. Four out of five classes reached maximum or near maximum scores with small standard deviations (SD) indicating that this knowledge was learned and retained by most students. These gains are displayed in Figure 1.

Also in Table 1, mean scores are displayed for the Sands Test of word spelling. In the fall, most students spelled none of the words correctly on the Sands Test. By May, substantial learning had taken place and all classes spelled a majority of words. Three of the five classes were close to maximum scores. The other two classes were able to spell on average three of the five words. These gains are displayed in Figure 2.

In Table 2 mean scores are shown for a second measure of letter-sound decoding on the Gates-MacGinitie test. This task contains double the items (30) of the previous letter-sound test to further assess alphabetic knowledge. In the fall, kindergarten classes recognized from a third to a half of the letter-sound items on this assessment. By May, all classes showed marked improvement and recognized from three-quarters to almost all the letter sounds. Of particular interest are the substantial gains by students in Teacher 1 and Teacher 3's classes. In Teacher 1's class, student performance was the lowest average score in September ($M = 9.19$) and rose to one of the highest by May ($M = 26.82$). These gains are shown in Figure 3.

Table 2 also reports mean spelling scores on the Morrison-McCall Spelling scale. In March kindergarteners were given this task as a pretest. Mean scores exceeded the expected grade-equivalent level (K), with GE scores ranging from 1.5 to 1.9 across classes. By May, all classes had improved in spelling. Grade equivalents exceeded expectations even further with levels ranging from means of 1.8 to 2.3. These gains are displayed in Figure 4.

In Figures 1, 2, 3, and 4, mean scores are displayed for test points in September and May by classroom teacher for letter-sound recognition and spelling. All classes in this study showed improvement.

Figure 1 shows that all kindergarten classes improved on letter-sound knowledge. Most notable in Figure 1 are the two kindergarten classes of Teachers 1 and 3 who showed the lowest mean scores in September but improved substantially by May.

Figure 2 shows that all kindergarten classes improved on word spelling. In the fall students were unable to write any words. By May all students could write some words with three classes close to the maximum score.

In Figure 3, the gains displayed for the Gates-MacGinitie subtest of letter-sounds show that students made great gains in learning this skill. Of note are the performances by students in Teacher 1 and Teacher 3 classes. These classes made especially large gains with scores rising from lower levels of performance to levels near the highest.

Figure 4 shows gains made in spelling on the Morrison McCall Spelling scale. All classes improved substantially in spelling. Again, of special note are student performances from Teacher 1 and Teacher 3's classes. Mean scores increased by more than half for Teacher 3's class.

In monthly logs, coaches reported that all classroom teachers had mastered the major competencies or skills that teachers were expected to master during the school year as a result of participating in the RRF program.

Coaches reported that all kindergarten classroom teachers had students with noted behavior problems that were disruptive to the implementation of lessons. In addition, two classrooms were noted for the other demands required of the class that restricted instructional time. In the two lowest performing classes, one class had many new students added throughout the year, as well as high attrition, so continuity of instruction was difficult. Also of note was that two of the lowest performing classes had students who were very learning disabled or who had special needs, making handwriting particularly challenging and retention of lesson content extremely difficult.

Nevertheless, in all five classes, students improved on letter-sound knowledge and spelling ability throughout the year. Coaches reported that all teachers tried hard and believed in the Spalding program or became increasingly committed to the program and coaching support during the year. In addition, all teachers were noted as having reached 75% to 100% mastery in pronouncing sounds, introducing new sounds, conducting dictation of words to students, teaching handwriting and spelling rules by the end of December, January, February, or March. Comments included how hard teachers tried, and even "exceptional" effort was noted for classrooms that had children with the most identifiable challenges.

In first grade:

Tables 3 and 4 report mean scores for first grade classes on measures of word decoding and spelling and show that performance was greater on May posttests than on fall pretests. These gains are depicted in Figures 5 and 6 for each measure.

In Table 3 mean scores are reported for first grader's word decoding. In the fall, classes performed at a beginning first grade level or below on word reading with mean GE scores ranging from K to 1.3. Word decoding improved from September to May across all classrooms. Final grade equivalent levels ranged from 1.2 to 2.5. These gains are shown in Figure 5.

In Table 4 mean scores are reported for first grader's spelling. In the fall, mean GEs ranged from 1.0 to 2.2 showing that most classes were at a mid-first grade level or above on spelling. By May all classrooms improved, some by almost a full grade level, with GEs ranging from 2.2 to 3.7. All classes exceeded grade level expectations in spelling by the end of first grade. These gains are displayed in Figure 6.

Figures 5 and 6 display impressive word decoding and spelling gains across classes for first graders. All classes in this study showed improvement from September to May.

In Figure 5, word-decoding gains show that many classes who were on grade level in the fall exceeded GE expectations in May. In addition, low performing classes in the fall improved to meet grade level expectations. Of particular note are the substantial gains for the low-performing classes of Teacher 5, Teacher 6 and Teacher 8.

Figure 6 shows that all first grade classes' spelling scores increased by at least an entire grade level from fall to spring, with gains ranging from 1.1 to 1.6 GE levels. All classes exceeded GE expectations for first grade by May, with GE scores ranging from 2.2 to 3.7. This included Classes 5, 8, and 9, the lowest performing classes, who also exceeded grade level expectations.

Coaches reported that all teachers reached 75% to 100% mastery of competencies and skills for delivering instruction to students by the end of December, January, February, or March. These competencies and skills included the ability to pronounce sounds, to introduce new sounds, to conduct dictation of words to students, and to teach handwriting and spelling rules. Many teachers were noted for their engagement and commitment to the Spalding program, their excellent mastery of instruction, their willingness to be creative, and their ability to handle challenges in administering the program consistently.

Behavior problems that were disruptive to the implementation of lessons were noted for three classes. One class was noted as having a large proportion of students with true behavior difficulties, which caused many interruptions to the instruction. Still, all first grade classes made substantial progress exceeding grade level expectations by the end of the year.

Two classrooms were noted for the other demands on the teachers, which restricted instructional time. However, these classroom teachers were noted for their commitment, hard work, increasing confidence and excellence, and all students in these classes made substantial progress during the year.

In second grade:

Tables 5 and 6 report mean scores for second grade classes on measures of word reading, comprehension, and spelling. Mean performance was greater on May posttests than on fall pretests. These gains are depicted in Figures 7, 8, and 9 for the three measures.

In Table 5, mean scores are reported for second grader's word decoding. Mean scores across classes ranged from GEs of 1.3 to 2.2 with only two classes scoring on grade level. Grade equivalent scores indicated that two of the six classes performed substantially below grade level in the fall (mean GEs = 1.3 and 1.5). By May, substantial improvement in decoding was made, with GEs ranging from 1.8 to 3.3. Four classes reached a second-grade level and one a third grade level. Although progress was made, the lowest performing class in September remained below a second grade level in May with a GE of 1.8. These gains are displayed in Figure 7.

Table 5 also shows second grade classes' mean performance on the Gates MacGinitie reading comprehension subtest. GEs revealed that five of the six classes were below a 2.0 grade level in the fall, with GEs ranging from 1.4 to 1.9. The exception performed at grade level. By May, scores increased for all classes with GEs ranging from 1.6 to 3.1. Four classes were on a second-grade level with GEs ranging from 2.0 to 2.7. One class exceeded grade level expectations with a GE score of 3.1. Although progress was made, the lowest performing class in the fall was still below grade level in May with a GE of 1.6, showing a GE gain of only 0.2. These scores are displayed in Figure 8.

Table 6 shows mean performance for second grade classes on spelling. In the fall, GEs ranged from 2.0 to 3.3. Four classes met grade level expectations with fall GEs ranging from 2.0 to 2.9. Two classes exceeded grade level expectations with GEs of 3.1 and 3.3. In May, all classes exceeded grade level expectations on spelling with GEs ranging from 3.5 to 4.7. The gains in scores by class are shown in Figure 9.

Figures 7, 8, and 9 display gains on word decoding, reading comprehension, and spelling across the second grade classes. All showed improvement from the fall to May.

In Figure 7, impressive word decoding gains across classes are displayed. Of particular interest are performance gains shown by students of Teachers 1, 2 and 4. These classes showed exceptional improvement in word reading, with gains of 0.9 to 1.1 GE levels.

Figure 8 displays average gains across second grade classes in reading comprehension. All classes in this study showed improvement from September to May. Of note is the exceptionally large mean gain of Teacher 2's class.

Figure 9 displays performance gains on the spelling test. Most notable are the classes of Teachers 3 and 4. These two classes made especially large gains from September to May on spelling. Figure 9 also shows that all of the second grade classes' scores increased by more than one grade level, with means ranging from 3.5 to 4.7 GE levels.

Coaches reported that most teachers reached 75% to 100% mastery of competencies and skills for delivering instruction to students. This included pronouncing sounds, introducing new sounds, conducting dictation of words to students, teaching handwriting and spelling rules by the end of

December, January, February, or March. The exceptions were two teachers whose proficiency in teaching spelling rules had not been achieved by March. However, both these teachers' classes made good progress in spelling.

Behavior problems from severe to minor were noted for three classes. Although these classes made gains in reading and spelling, two of these classes were among the lowest performing across all measures. In addition, disruptions from outside of the classroom were noted for these two classes. Also, new students were admitted to these two classes making continuity of instruction a challenge. Still, all second grade classes made substantial gains. All classes performed at least at a second grade-equivalent level on the decoding and comprehension measures except for one class, and all classes exceeded grade level expectations on the spelling measure.

Conclusion

Substantial growth was found across all measures for all kindergarten, first and second grade classes. The instructional program that involved RRF teacher coaching combined with the Spalding method of explicit phonics instruction improved students' alphabetic knowledge, word reading, reading comprehension, and spelling performance. Regardless of the type of classroom (i.e., general education, ICT, or self-contained) all classes improved significantly from September to May on literacy skills.

A combination of teacher effort and commitment, coaching support by well-trained expert providers, and fidelity to the treatment of the Spalding method are all factors that appear to have strongly contributed to the improvement of literacy skills seen across all types of classes with many different types of learners.

Table 1.   Kindergarten classes letter/sound and words

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses for kindergarten classes

Teacher[e] Number	School	N[b] (72)	Class Type[c]	% Free Lunch[a]	September[d] Sands Test		May[d] Sands Test	
					Letter-Sounds (max=15) M SD	Spell Words (max=5) M SD	Letter-Sounds (max=15) M SD	Spell Words (max=5) M SD
1	277	15	Gen	93%	8.88 (3.2)	0.31 (0.6)	15.00 (0.0)	4.41 (0.7)
2	32	8	ICT	50%	10.44 (4.2)	0.44 (1.0)	15.00 (0.0)	4.00 (1.8)
3	208	22	Gen	79%	8.32 (4.5)	0.77 (3.0)	14.41 (1.9)	4.18 (1.6)
4	117	18	ICT	72%	10.33 (3.6)	0.44 (1.1)	14.58 (1.4)	3.68 (1.9)
5	117	9	SC	72%	10.00 (5.5)	0.56 (1.7)	13.11 (5.0)	3.67 (1.9)

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website.

[b] Only students who took the pretests and posttests were included in the analyses.

[c] Class types: Gen=General Education, ICT=Inclusive Collaborative Teaching, SC=Self Contained.

[d] Mean (M) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (SD) = average difference between individual scores and the mean score. Max = total number of items.

[e] Teacher numbers in the first column correspond to teacher numbers in Figures 1 and 2.

Table 2.   Kindergarten classes letter/sound and spelling

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses, and mean grade equivalent (GE) scores for kindergarten classes

Teacher[g] Number	School	N[b] (72)	Class Type[e]	% Free Lunch[a]	Gates Decode[f] September Letters/Sounds[c] RS (max=30) M	SD	May M	SD	Morrison McCall[f] March Spelling Words[d] RS (max=10) M	SD	GE	May RS (max=10) M	SD	GE
1	277	15	Gen	93%	9.19	(4.7)	26.82	(2.1)	3.88	(3.1)	1.8	6.69	(3.2)	2.1
2	32	8	ICT	50%	16.89	(6.4)	28.56	(2.2)	4.78	(5.7)	1.9	8.67	(5.7)	2.3
3	208	22	Gen	79%	13.18	(6.4)	26.91	(5.1)	3.32	(2.3)	1.7	7.18	(3.1)	2.1
4	117	18	ICT	72%	16.17	(7.4)	24.90	(4.4)	4.39	(2.8)	1.8	5.95	(3.6)	2.0
5	117	9	SC	72%	11.00	(6.2)	22.78	(6.1)	1.56	(2.0)	1.5	3.78	(3.4)	1.8

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website.

[b] Only students who took the pretests and posttests were included in the analyses.

[c] Only raw scores (RS), not grade equivalents, are available for the Gates test of decoding letters and letter sounds at the Level K, Form PR.

[d] Morrison-McCall Spelling test, List # 6 was administered in March, List # 1 in May. The only exception was Teacher Class 1 where List # 1 was administered in March and May.

[e] Class types: Gen=General Education, ICT=Inclusive Collaborative Teaching, SC=Self Contained

[f] Mean (M) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (SD) = average difference between individual scores and the mean score. Max = total number of items. GE = mean grade equivalent score.

[g] Teacher numbers in the first column correspond to teacher numbers in Figures 3 and 4.

Table 3.   First grade classes word decoding

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses, and mean grade equivalent (GE) scores for first grade

Teacher[e] Number	School Number	N[b] (201)	Class Type[d]	% Free Lunch[a]	Gates Decode Words[c] September[f] RS (max=43) M SD	GE	Gates Decode Words[c] May[f] RS (max=43) M SD	GE
1	277	25	Gen	93%	15.96 (7.5)	1.0	38.20 (5.8)	2.5
2	277	26	ICT	93%	14.69 (10.4)	1.0	30.81 (8.5)	1.9
3	208	21	Gen	79%	20.67 (11.1)	1.3	31.00 (8.9)	1.9
4	208	28	Gen	79%	19.22 (10.7)	1.2	33.54 (8.7)	2.1
5	14	19	Bil/ICT	76%	6.62 (5.9)	K	26.05 (10.8)	1.6
6	14	25	Gen	76%	9.96 (7.2)	K	35.16 (6.7)	2.2
7	117	30	Gen	83%	20.33 (8.5)	1.3	37.27 (6.1)	2.3
8	57	20	Gen	83%	10.15 (4.6)	K	29.35 (9.5)	1.8
9	57	7	SC 1st-2nd	83%	7.33 (4.0)	K	19.33 (11.0)	1.2

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website

[b] Only students who took the pretests and posttests were included in the data analyses. RS = raw scores.

[c] The Gates-MacGinitie word decoding test, First Grade, Level 1, Forms S was administered in September and May.

[d] Class types: Gen=General Education, ICT=Inclusive Collaborative Teaching, BIL=Bilingual, SC=Self Contained

[e] Teacher numbers in the first column correspond to teacher numbers in Figure 5.

[f] Mean (M) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (SD) = average difference between individual scores and the mean score. RS = raw scores. Max = total number of items. GE = mean grade equivalent score.

[g] Both 1st and 2nd graders in the SC class of Teacher 9 received the same instruction and completed the first grade assessment measures.

Table 4.   First grade classes spelling

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses, and mean grade equivalent (GE) scores for first grade

Teacher[e] Number	School Number	N[b] (201)	Class Type[d]	% Free Lunch[a]	Morrison McCall Spelling September[c] RS (max=30) M SD	GE	Morrison McCall Spelling May[c] RS (max=30) M SD	GE
1	277	25	Gen	93%	4.58 (2.8)	1.9	18.08 (5.7)	3.5
2	277	26	ICT	93%	3.62 (3.0)	1.8	14.81 (6.2)	3.1
3	208	21	Gen	79%	7.62 (5.5)	2.2	17.10 (6.7)	3.4
4	208	28	Gen	79%	6.71 (6.0)	2.1	17.43 (6.7)	3.4
5	14	19	Bil/ICT	76%	1.86 (3.1)	1.5	14.58 (5.5)	3.1
6	14	25	Gen	76%	2.80 (3.1)	1.7	14.40 (5.1)	2.9
7	117	30	Gen	83%	6.71 (4.4)	2.1	19.13 (5.1)	3.7
8	57	20	Gen	83%	1.75 (2.7)	1.5	11.75 (5.7)	2.6
9	57	7	SC 1st-2nd	83%	0.00 (0.0)	1.0	8.00 (7.2)	2.2

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website

[b] Only students who took the pretest and posttest were included in the data analyses.

[c] Morrison-McCall Spelling test, List # 1, was administered in September and in May.

[d] Class types: Gen=General Education, ICT=Inclusive Collaborative Teaching, BIL=Bilingual, SC=Self Contained

[e] Teacher numbers in the first column correspond to teacher numbers in Figure 6.

[f] Mean (M) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (SD) = average difference between individual scores and the mean score. RS = raw scores. Max = total number of items. GE = mean grade equivalent score.

[g] Both 1st and 2nd graders in the SC class of Teacher 9 received the same instruction and completed the first grade assessment measures.

Table 5. Second grade classes word decoding and comprehension

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses, and mean grade equivalent (GE) scores for second grade classes

Teacher[e] Number	School N[b] (118)	Class N	Class Type[d]	% Free Lunch[a]	Gates Tests September[c]				Gates Tests May[c]			
					Decode[f] RS (max=43) M (SD)	GE	Comprehension[f] RS (max=39) M (SD)	GE	Decode[f] RS (max=43) M (SD)	GE	Comprehension[f] RS (max=39) M (SD)	GE
1	452	25	ICT	9	31.45 (11.0)	2.2	27.57 (10.8)	2.2	39.54 (6.8)	3.3	33.71 (5.9)	3.1
2	32	14	ICT	50	23.43 (14.2)	1.8	18.43 (10.4)	1.6	38.07 (6.3)	2.9	32.00 (5.6)	2.7
3	208	25	Gen	79	27.40 (10.5)	2.0	23.75 (7.0)	1.9	34.44 (8.1)	2.4	30.84 (5.1)	2.5
4	14	28	Gen	76	24.00 (10.2)	1.8	22.24 (9.4)	1.8	37.04 (6.7)	2.7	29.21 (7.0)	2.3
5	57	6	SC	83	14.86 (8.7)	1.3	14.50 (5.9)	1.4	23.83 (8.7)	1.8	17.67 (4.7)	1.6
6	57	20	Gen	83	19.00 (9.9)	1.5	16.60 (6.3)	1.5	31.75 (9.8)	2.3	26.40 (5.8)	2.0

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website

[b] Only students who took the spelling test in both September and May were included in the data analyses.

[c] The Gates-MacGinitie test, Second Grade Level 2, Form S was administered in September, Form T was administered in May.

[d] Class types are: Gen=General Education, ICT=Inclusive Collaborative Teaching, SC=Self Contained

[e] Teacher numbers in the first column correspond to teacher numbers in Figures 7 and 8.

[f] Mean (*M*) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (*SD*) = average difference between individual scores and the mean score. RS = raw scores. Max = maximum number of items. GE = mean grade equivalent score.

Table 6.   Second grade classes spelling

Schools, teachers, characteristics of classes, mean performance of students, standard deviations in parentheses, and mean grade equivalent (GE) scores for second grade classes

Teacher[e] Number	School[b] (118)	N[b]	Class Type[d]	% Free Lunch[a]	Morrison McCall Spelling Test[c]				
					September[f]		May[f]		
					RS (max=40)	GE	RS (max=40)		GE
					M SD		M SD		
1	452	25	ICT	9	14.83 (7.9)	3.1	25.46 (8.1)		4.7
2	32	14	ICT	50	15.79 (8.3)	3.3	23.57 (8.5)		4.5
3	208	25	Gen	79	14.04 (6.3)	2.9	24.44 (7.4)		4.5
4	14	28	Gen	76	13.62 (6.2)	2.9	25.39 (7.6)		4.7
5	57	6	SC	83	5.71 (5.7)	2.0	18.00 (6.1)		3.5
6	57	20	Gen	83	8.90 (3.5)	2.3	19.20 (5.8)		3.7

[a] Percentages of students eligible for free lunch were obtained using the 2014-15 School Quality Guide found on the New York City Department of Education's School Enrollment and Demographic Data section on their website

[b] Only students who took the spelling test in both September and May were included in the data analyses.

[c] Morrison-McCall Spelling test, List # 1 was administered in September and in May.

[d] Class types are: Gen=General Education, ICT=Inclusive Collaborative Teaching, SC=Self Contained

[e] Teacher numbers in the first column correspond to teacher numbers in Figure 9.

[f] Mean (*M*) = sum of scores of students divided by the number of students to reveal average performance of each class. Standard Deviation (*SD*) = average difference between individual scores and the mean score. RS = raw scores. GE = mean grade equivalent scores.

Figure 1. Mean Letter-Sounds Read Correctly in September and May on Sands Word Reading Test by Kindergarten Classes (15 max)
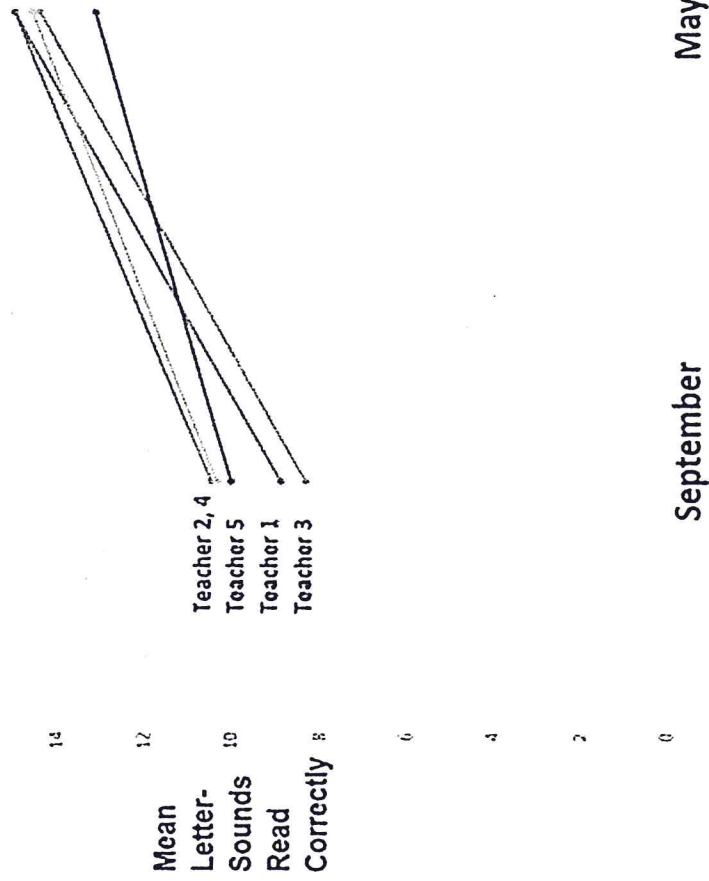
Mean Letter-Sounds Read Correctly

14
12
10
8
6
4
2
0

Teacher 2, 4
Teacher 5
Teacher 1
Teacher 3

September   May

Figure 2. Mean Words Spelled Correctly in September and May on Sands Word Reading Test by Kindergarten Classes (5 max)

Mean
Words
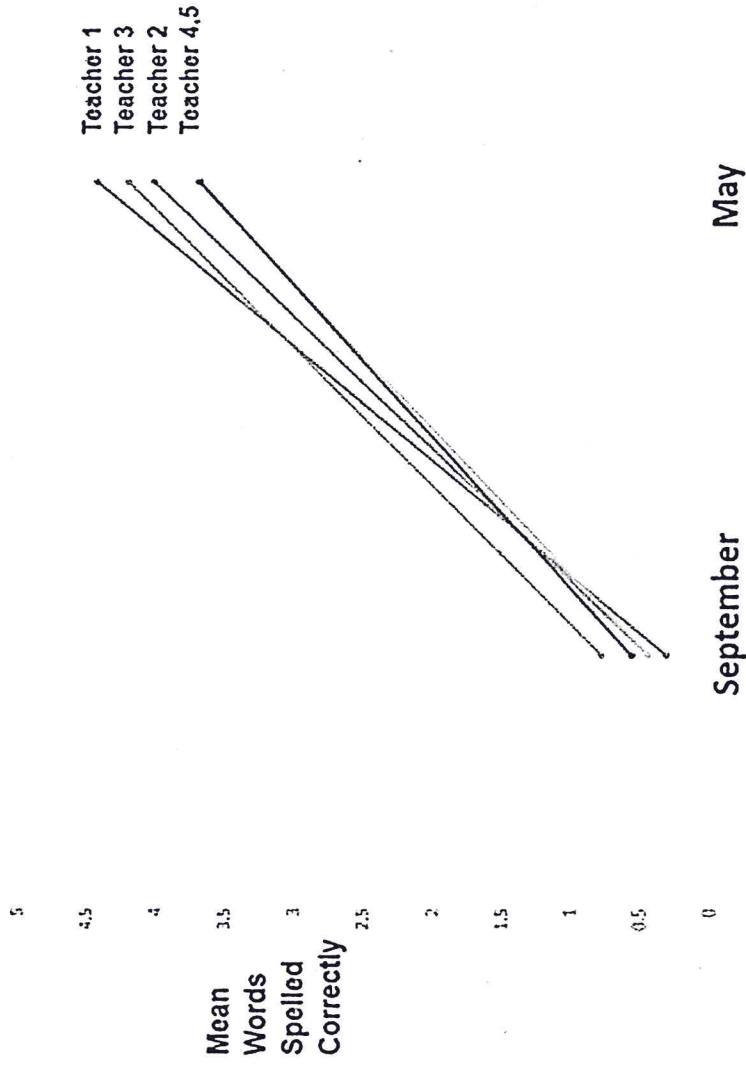Spelled
Correctly

5

4.5

4

3.5

3

2.5

2

1.5

1

0.5

0

September

May

Teacher 1
Teacher 3
Teacher 2
Teacher 4.5

Figure 3. Mean Letter-Sounds in September and May on Gates MacGinitie Decoding Test by Kindergarten Classes (30 max)

Mean
Letter-
Sounds
Correct

30

25

20

15

10

5

0

Teacher 2, 4

Teacher 3

Teacher 5
Teacher 1

September

May
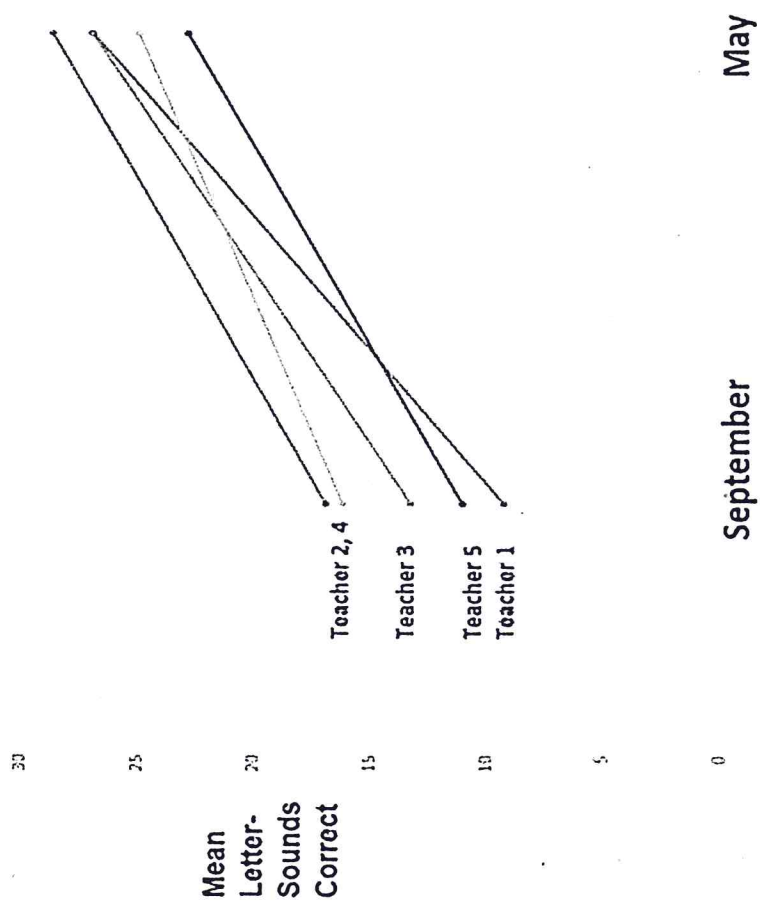
Figure 4. Mean Words Spelled Correctly in March and May on Morrison McCall Spelling Test by Kindergarten Classes (10 max)

Mean
Words
Correct

10
9
8
7
6
5
4
3
2
1
0

Teacher 2
Teacher 4
Teacher 1
Teacher 3
Teacher 5

September   May

Grade
Equivalent
2.3

2.1

1.9

1.7

1.3
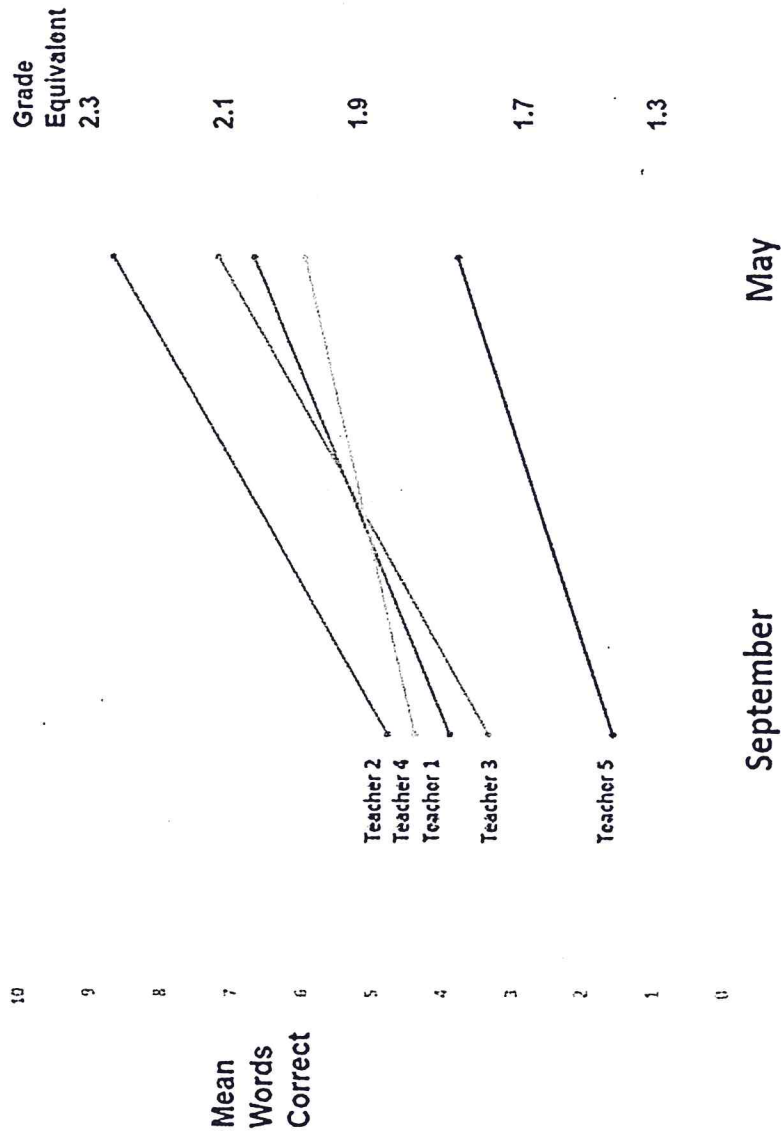
Figure 5. Mean Words Decoded Correctly in September and May on the Gates MacGinitie Decoding Test by First Grade Classes (43 max)
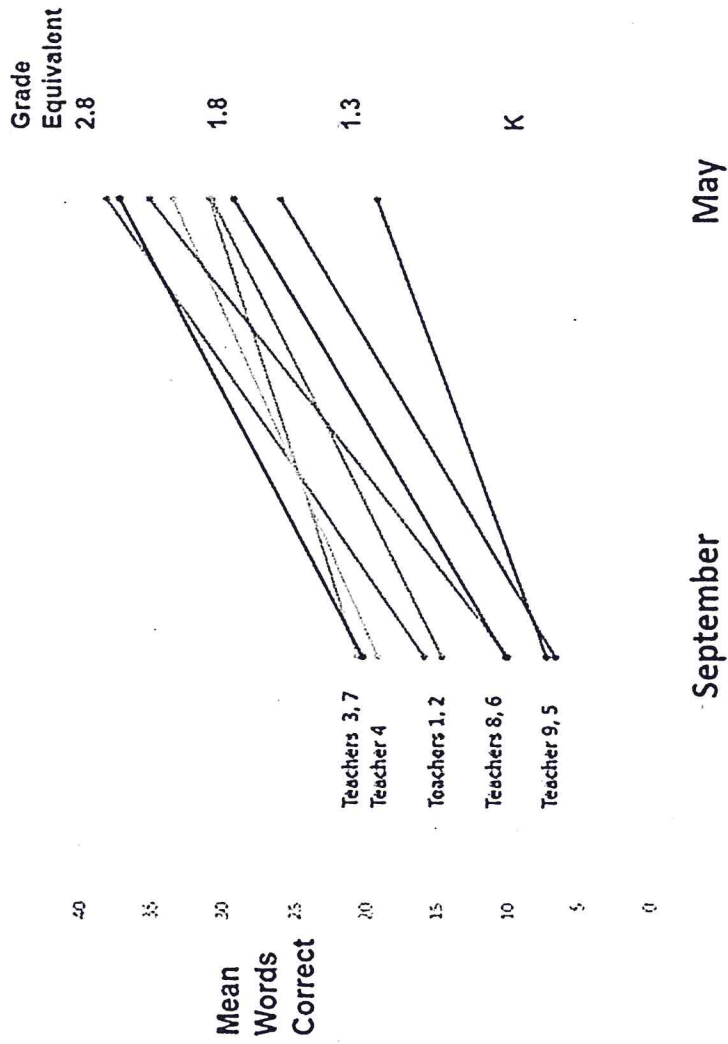
Figure 6. Mean Words Spelled Correctly in September and May on Morrison McCall Spelling Test by First Grade Classes (30 max)
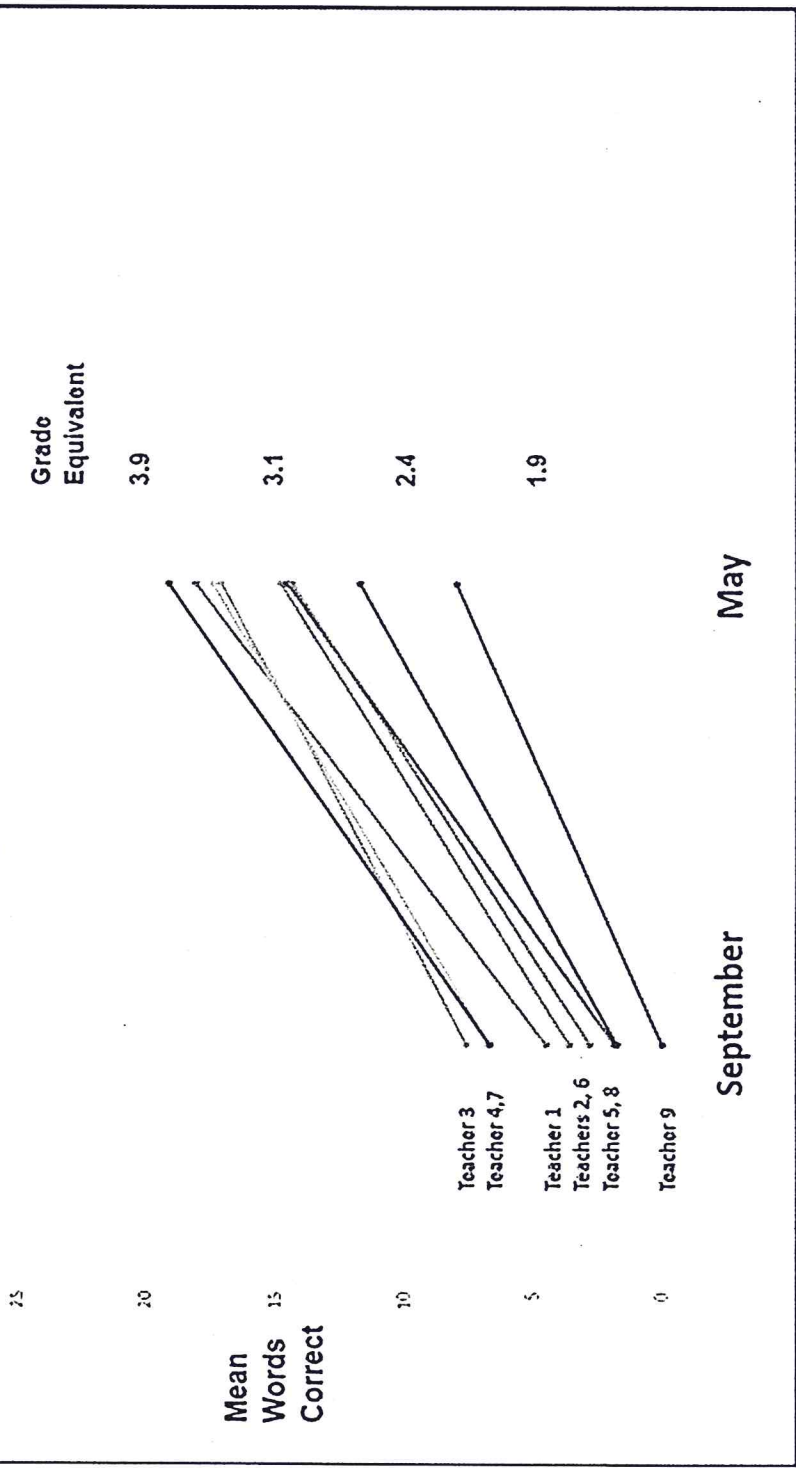
Figure 7. Mean Words Decoded Correctly in September and May on the Gates MacGinitie Decoding Test by Second Grade Classes (43 max)
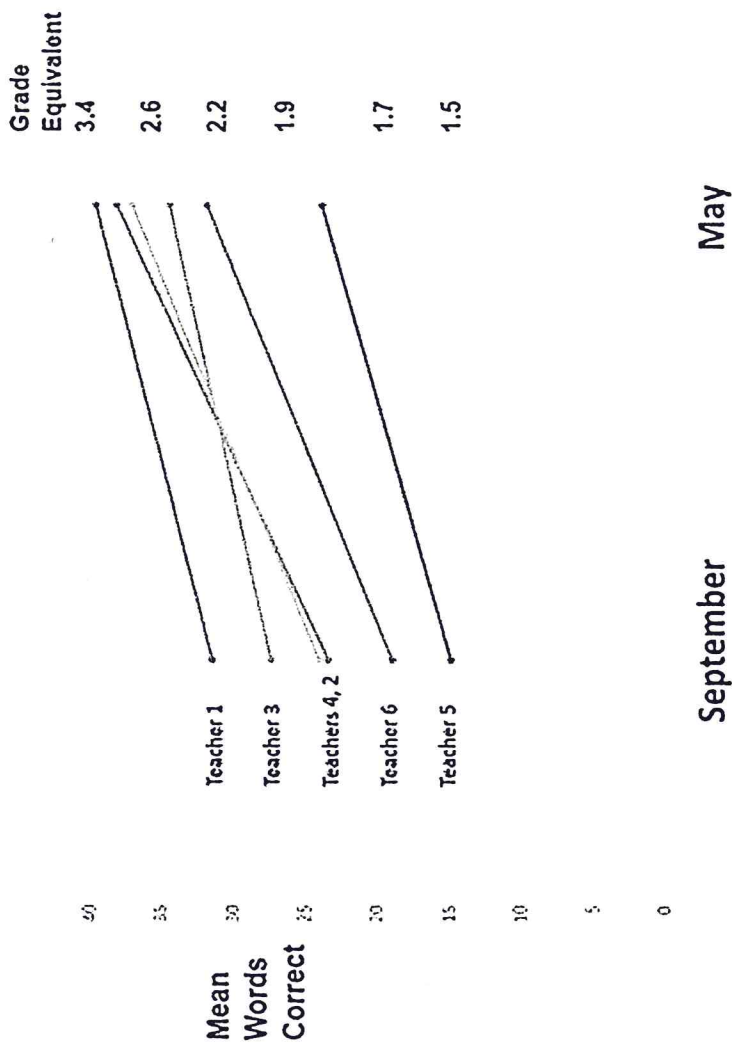
Figure 8. Mean Scores in September and May on the Gates MacGinitie
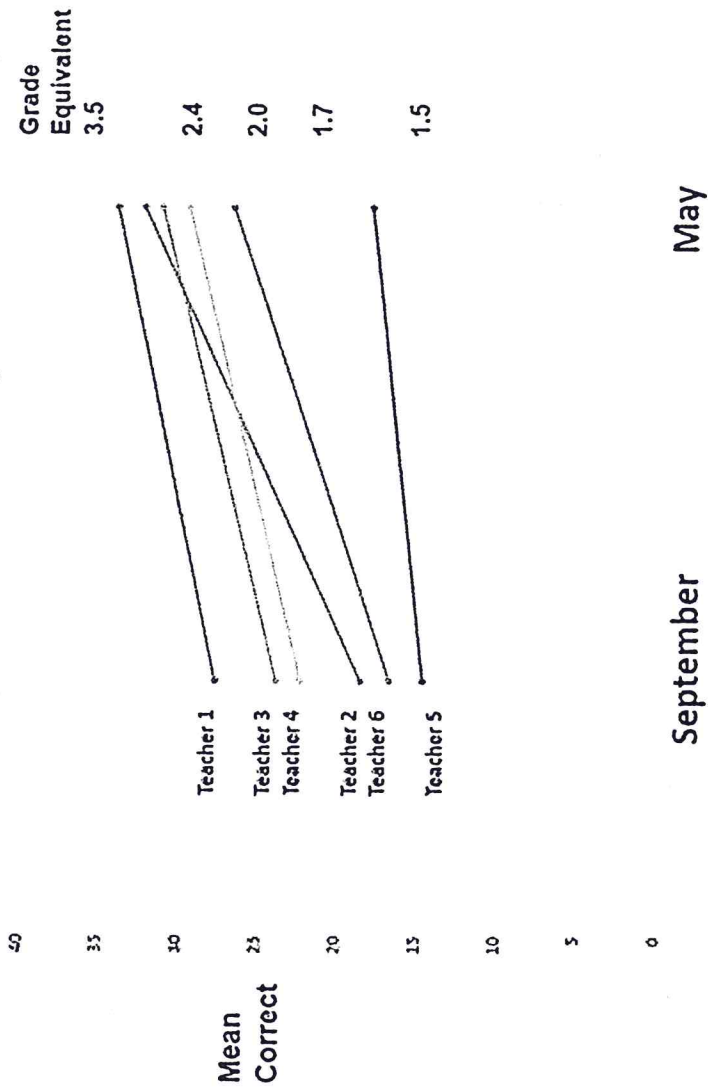Reading Comprehension Test by Second Grade Classes (43 max)

Figure 9. Mean Words Spelled Correctly in September and May on Morrison McCall Spelling Test by Second Grade Classes (40 max)