

December 2016

This research brief was prepared by American Institutes for Research (AIR) by Martyna Citkowicz, Ph.D.; Melissa Brown-Sims; Ryan Williams, Ph.D.; and Dean Gerdeman, Ph.D.

AIR is one of the world's largest behavioral and social science research and evaluation organizations. AIR's mission is to conduct and apply the best behavioral and social science research and evaluation towards improving people's lives, with a special emphasis on the disadvantaged.

Iowa's Teacher Leadership and Compensation Program

Evaluation of the Teacher Leadership and Compensation Program

The Iowa Teacher Leadership and Compensation (TLC) program was launched in the 2014–15 academic year with the following five goals:¹

- (1) Attract able and promising new teachers by offering competitive starting salaries and offering short-term and long-term professional development and leadership opportunities.
- (2) Retain effective teachers by providing enhanced career opportunities.
- (3) Promote collaboration by developing and supporting opportunities for teachers in schools and school districts statewide to learn from each other.
- (4) Reward professional growth and effective teaching by providing pathways for career opportunities that come with increased leadership responsibilities and involve increased compensation.
- (5) Improve student achievement by strengthening instruction.

The TLC program was rolled out in three successive district cohorts, each covering approximately one third of Iowa's students:

- Cohort 1 in 2014–15 ($n = 39$ districts)
- Cohort 2 in 2015–16 ($n = 76$ districts)
- Cohort 3 in 2016–17 ($n = 221$ districts)

The Iowa Department of Education (DE) contracted American Institutes for Research (AIR) to evaluate TLC in June 2015. The evaluation was designed to inform the DE about the progress of the TLC initiative and its effectiveness in meeting its goals.

This brief highlights interim findings from early implementation of the TLC program to provide the DE with initial feedback on TLC progress toward intended goals. The brief includes findings on implementation progress from teacher and administrator surveys, interviews and focus groups and includes results of initial analyses of student achievement in Cohort 1 districts in 2014–15 compared with districts that had not yet implemented TLC (Cohorts 2 and 3). These findings are considered interim because the evaluation is in progress, with additional data on TLC implementation to collect and analyze over the next two years.

¹ These are available on the Iowa Department of Education's website: <https://www.educateiowa.gov/teacher-leadership-and-compensation-system>

Evaluation of TLC Implementation

To examine TLC implementation in Cohort 1 and Cohort 2 districts in 2015–16, we administered surveys and conducted focus groups and interviews in spring 2016.

Surveys. We administered statewide online surveys to Iowa teachers and school and district administrators to obtain feedback on changes that accompanied the TLC program implementation. The survey included items related to four areas of change in the TLC program: teacher leadership roles; professional development and supports for teachers; opportunities for teacher collaboration; perceived outcomes of TLC implementation.

Overall, 30% of teachers ($n = 10,746$), 50% of school administrators ($n = 738$), and 22% of district administrators ($n = 316$) in the DE's Basic Educational Data Survey database completed the survey.² To reduce a large number of survey items to a smaller set of key constructs, we constructed scale scores by combining related survey items and calculated response percentages for the scale scores.³ (See Box 1.)

Focus groups and interviews. We invited a randomly selected group of Cohort 1 and Cohort 2 teacher leaders across 10 randomly selected districts to participate in 90-minute focus groups on TLC program implementation. Ninety-three teacher leaders (39 from Cohort 1 and 54 from Cohort 2) participated, including those in lead teacher, mentor teacher, model teacher, instructional coach, and curriculum or professional development leader roles. We also interviewed 12 Area Education Agency (AEA) staff⁴ across all nine AEAs in the state, as well as 10 superintendents or assistant superintendents from the selected districts. Focus groups utilized iClicker software, which allowed for quick and anonymous polling of the respondents. We analyzed iClicker response data, transcripts for patterns, themes, and categories to determine the most important findings and key similarities and differences across the focus group and interview responses.

Box 1. Survey Scale Scores. Scale scores indicate the degree to which a measured construct is present (for example, the degree to which respondents agree to a set of statements about the utility of supports provided). Higher scores indicate a more positive perception, whereas lower scores indicate a more negative perception, generally speaking. We categorized these scores along the original response options for each construct (e.g., *disagree strongly*, *disagree somewhat*, *agree somewhat*, or *agree strongly*), where the lowest scale scores were categorized in the lowest response categories and the highest scale scores were in the highest response categories. We then calculated percentages of respondents in each category in the scale to highlight the typical responses from surveyed teachers and administrators. For example, when we asked a series of questions about the presence of opportunities to advance into leadership roles, the 75% of surveyed teachers fell in either the *agree strongly* or *agree somewhat* categories, indicating opportunities to advance were available and attainable in their school or district. We calculated overall percentages, as well as percentages for Cohorts 1, 2, and 3, separately. For findings presented in this brief, we tested for cohort differences in the extreme category percentages, such as *agree strongly*.

² We examined the representativeness of our sample along years of experience and degree earned to examine sample representativeness using a raking technique (Battaglia, Hoaglin, & Frankel, 2009).

³ We used the Rasch rating scale model (Wright & Masters, 1982), using Winsteps® (Linacre, 2015) for these analyses.

⁴ AEA staff included chief administrators, assistant chief administrators, regional administrators, and executive directors.

Interim Findings on TLC Program Implementation

This section highlights initial findings about implementation from the perspectives of respondents. When possible, responses from TLC districts, including Cohort 1 (2014–15) and Cohort 2 (2015–16), were compared to responses from Cohort 3 (2016–17) districts, which had not yet launched the program at the time of the data collection. These cohorts allow comparisons based on the presence or absence of TLC (Cohorts 1 and 2 compared to the Cohort 3) and on the level of maturity of the TLC program (Cohort 1 compared to Cohort 2).

When possible, responses from TLC districts, including Cohort 1 (2014–15) and Cohort 2 (2015–16), were compared to responses from Cohort 3 (2016–17) districts, which had not yet launched the program at the time of the data collection.

The results from the survey, focus group, and interview analyses are correlational and descriptive in nature. They do not provide evidence about the effects of TLC in a causal framework. Differences in responses across cohorts could be due to preexisting differences among districts and respondents.

Respondents from TLC districts reported greater familiarity, availability, and perceived effectiveness of teacher leadership roles, especially among Cohort 1 respondents.

“My understanding is that we are working to help teachers be reflective practitioners and that a variety of different positions within the TLC grant are there to help facilitate the way that we implemented that.”

—Cohort 2 Teacher Leader

Familiarity with teacher leadership roles. Respondents from TLC districts were more familiar with teacher leadership roles. Survey respondents in Cohort 1 (57% of teachers and 85% of administrators) and Cohort 2 (52% of teachers and 80% of administrators) were significantly more likely to indicate that they are *very familiar* with teacher leadership roles than respondents in Cohort 3 (22% of teachers and 50% of administrators). Cohort 1 respondents were significantly more likely to be *very familiar* with teacher leadership roles than respondents in Cohort 2. Despite the greater familiarity in TLC districts, more than 40% of teacher

respondents in these districts still were not *very familiar* with teacher leadership roles, and teachers reported lower familiarity than administrators.

Availability of teacher leadership roles. Survey respondents from TLC districts reported greater availability of several teacher leadership roles compared to non-TLC districts, particularly lead teachers, model teachers, and instructional coaches (Figures 1 and 2). Administrators were more likely than teachers to report the presence of teacher leader roles. All focus group and interview respondents from TLC district and AEA administrators reported that instructional coaches were available in their districts. Moreover, respondents from 5 of the 10 focus group districts also reported the availability of mentor teacher and mentor coach roles.

Cohort 1 and Cohort 2 teachers and administrators reported significantly higher levels of familiarity with new teacher leadership roles compared to teachers and administrators from Cohort 3.

More teachers in TLC districts reported holding teacher leadership roles compared to non-TLC districts, but the differences were small. Overall, 27% of surveyed teachers indicated that they held a teacher leadership role; 26% in Cohort 1 indicated this, with 29% in Cohort 2 and 23% in Cohort 3.

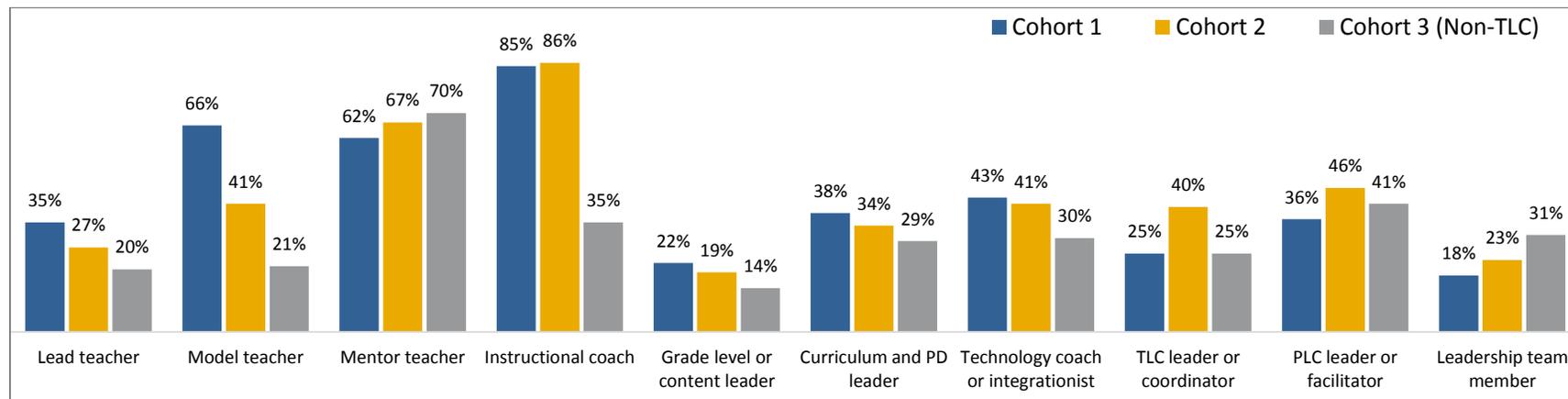
Opportunities to assume teacher leadership roles. No significant differences were found across surveyed teachers in TLC and non-TLC districts regarding opportunities to assume teacher leadership roles. When asked a series of questions about the presence of opportunities to advance into leadership roles, scale scores for 75% of surveyed teachers fell in the *agree strongly* or *agree somewhat* categories, indicating that these opportunities to advance were perceived as available and attainable in their school or district. Among the teacher leaders who participated in focus groups, all of whom worked in TLC districts, the large majority (81 out of 93) *agreed or agreed strongly* that they had personally participated in targeted professional development opportunities to build their skills as teacher leaders.

“One of the biggest things is that there’s a barrier between the coaches and the administration. So we are not evaluative. And we are simply there to help teachers. It’s had a positive impact because teachers can come with us. They come to us and ask us about anything that has to do with teaching.”

—Cohort 1 Teacher Leader

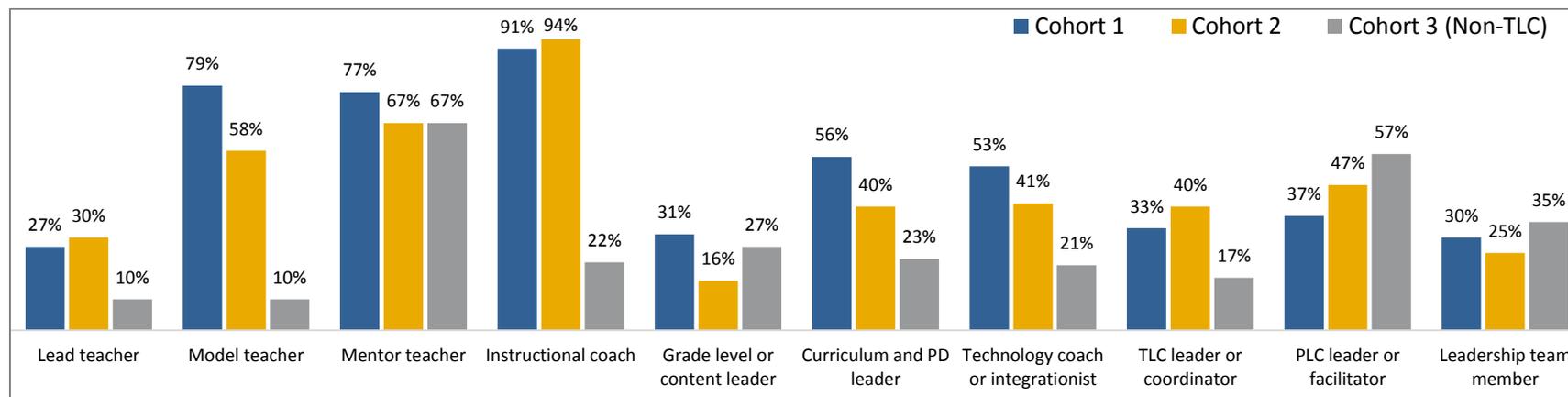
Effectiveness of teacher leadership roles. Most surveyed teachers and administrators perceived teacher leadership roles in their schools as effective. Based on a series of questions about the effectiveness of these roles, scores from a majority of respondents (87% of teachers and 96% of administrators) fell in the *agree somewhat* or *agree strongly* categories, indicating that these roles were perceived as effective in increasing support for teachers and improving student achievement through improved instruction. Scores from respondents in Cohort 1 (42% of teachers and 69% of administrators) and Cohort 2 (42% of teachers and 55% of administrators) were significantly more likely to fall in the *agree strongly* category compared to those in Cohort 3 (27% of teachers and 10% of administrators), indicating that early adopting TLC cohorts were more likely to view the teacher leadership roles as effective.

Figure 1. Percentage of Surveyed Teachers Who Reported Available Teacher Leadership Roles in Their School or District



Note: Sample sizes for surveyed teachers and teacher leaders: $n_{c1} = 3,493$, $n_{c2} = 3,210$, $n_{c3} = 1,118$. Response options selected by less than 10% of respondents, on average, and response options *Other* and *Do not know* were omitted.

Figure 2. Percentage of Surveyed Administrators Who Reported Available Teacher Leadership Roles in Their School or District



Note: Sample sizes for surveyed administrators: $n_{c1} = 287$, $n_{c2} = 299$, $n_{c3} = 118$. Response options selected by less than 10% of respondents, on average, and response options *Other* and *Do not know* were omitted.

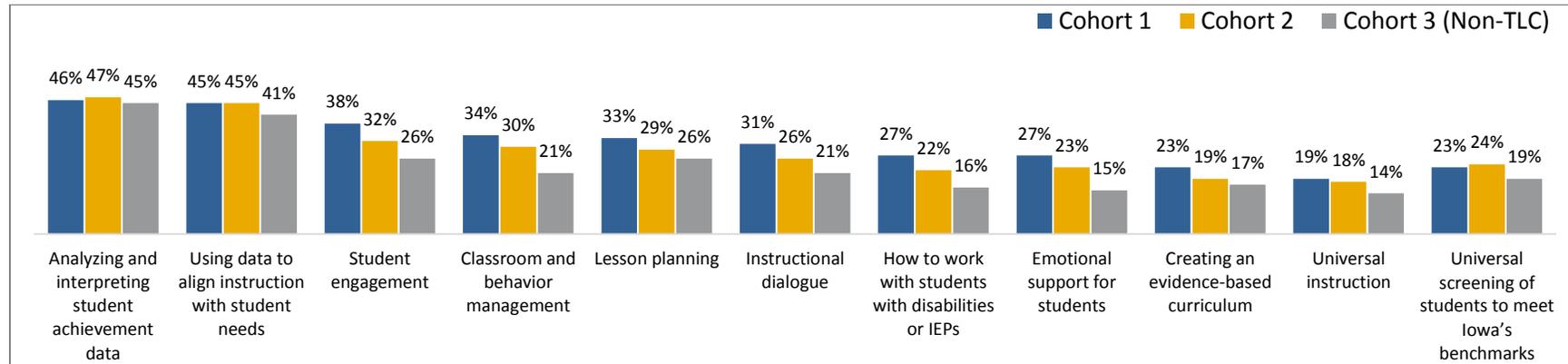
Respondents from TLC districts perceived greater availability, utility, and quality of supports for teachers and teacher leaders.

Available school- and district-level professional development supports. Across all three TLC cohorts, 99% of surveyed teachers indicated that they participated in some kind of professional development in the 2015–16 school year. Teachers in Cohort 1 (78% of teachers) were significantly more likely to indicate that professional development support was offered at the school and district levels than teachers in Cohort 2 (71% of teachers) or Cohort 3 (70% of teachers). According to interviews with 5 out of 12 AEA staff, their AEAs offer the Journey to Excellence training regionally, which is geared toward mentoring new teachers.

Teachers and school administrators across all three cohorts reported professional development offerings and supports that covered a variety of topics (Figures 3 and 4). Teachers in TLC districts more frequently reported most of the professional development offerings, with teachers in Cohort 1 more often reporting the offerings than teachers in Cohort 2. In nearly all cases, a higher percentage of administrators reported professional development offerings and supports compared to teachers. In addition, differences between TLC and non-TLC districts in reported available supports were not as consistent among administrators as they were among teachers.

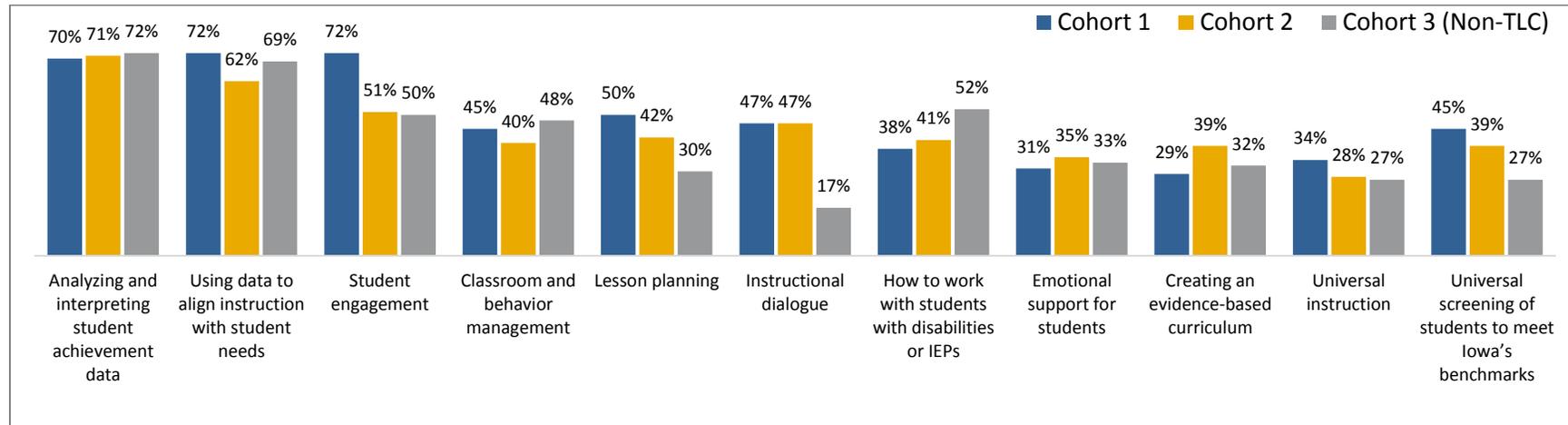
Utility and quality of available teacher supports. Surveyed TLC teachers were more likely to indicate that professional development supports provided were useful and of high quality compared to non-TLC teachers. Surveyed teachers and administrators responded to two sets of items: one about the utility of available supports and the other about the quality of available supports. Survey scale scores from most respondents across cohorts fell in the *agree somewhat* or *agree strongly* categories for their perceptions about the utility (80%) and quality (79%) of the supports provided for improving instruction. However, utility scale scores for teachers in Cohort 1 (25% of teachers) and Cohort 2 (26% of teachers) were significantly more likely to fall in the *agree strongly* category, indicating the provided supports were perceived as useful, compared to teachers in Cohort 3 (17% of teachers). Similarly, quality scores for teachers in Cohort 1 (26% of teachers) and Cohort 2 (27% of teachers) were significantly more likely to fall in the *agree strongly* category, indicating that the supports offered were perceived to be of high quality, compared to teachers in Cohort 3 (18% of teachers).

Figure 3. Percentage of Surveyed Teachers Who Reported* the Following Teacher Supports Offered in Their School or District



Note: Sample sizes for surveyed teachers: $n_{c1} = 3,495$, $n_{c2} = 3,204$, $n_{c3} = 3,622$; *Omitted response options include *Other* and *Do not know*.

Figure 4. Percentage of Surveyed Administrators Who Reported* the Following Teacher Supports Offered in Their School or District



Note: Sample sizes for surveyed administrators: $n_{c1} = 286$, $n_{c2} = 299$, $n_{c3} = 464$; *Omitted response options include *Other* and *Do not know*.

Respondents from TLC districts reported greater frequency of collaboration and satisfaction with opportunities for teachers to engage in collaboration.

Frequency of teacher collaboration. Compared to non-TLC teachers and administrators, surveyed TLC teachers and administrators more frequently indicated participation in weekly collaboration activities (Figures 5 and 6). Teachers and administrators generally responded similarly about their perceptions of collaboration frequency.

Nearly all of the teacher leaders participating in focus groups (98%) *agreed* or *strongly agreed* that the TLC program provided teachers with opportunities to engage in high-quality collaboration with their peers. The most frequently cited mechanism was professional learning communities.

Satisfaction with opportunities to engage in teacher collaboration. Compared to non-TLC teacher survey respondents, TLC teachers were more likely to be satisfied with their collaboration opportunities. When asked a series of questions about their satisfaction with the productivity and facilitation of teacher collaboration, survey scale scores for 82% of all surveyed teachers fell in either the *agree strongly* or the *agree somewhat* categories, indicating that teacher collaboration was perceived to be both productive and well facilitated. The percentage of scale scores in the *agree strongly* category for teachers in Cohort 1 (30% of teachers) and Cohort 2 (30% of teachers) was larger than Cohort 3 (21% of teachers), indicating significantly greater perceived satisfaction with collaboration opportunities in the early adopting TLC cohorts.

Most respondents in TLC Cohorts 1 and 2 reported that TLC had a positive impact on teaching instruction and changed their professional structure.

“We don’t know yet. . . . It takes four to five years . . . from what research says to see impact. I can tell you looking at the high school test scores we didn’t have many kids make growth. We had a lot of proficiency, but we didn’t have a lot of kids make growth in certain areas . . . I think it’s having an impact . . . it’s going to trickle down to the kids. Are we seeing fantastic numbers yet? No. Will we? I’m hoping so.”

—Cohort 1 Teacher Leader

TLC’s impact on instruction. Surveyed TLC teachers and administrators perceived that TLC is improving instruction, with similar perceptions for respondents in Cohorts 1 and 2. When asked a series of questions about the effectiveness of the TLC program, scale scores for 77% of teachers and 87% of administrators fell in either the *agree somewhat* or the *agree strongly* categories.

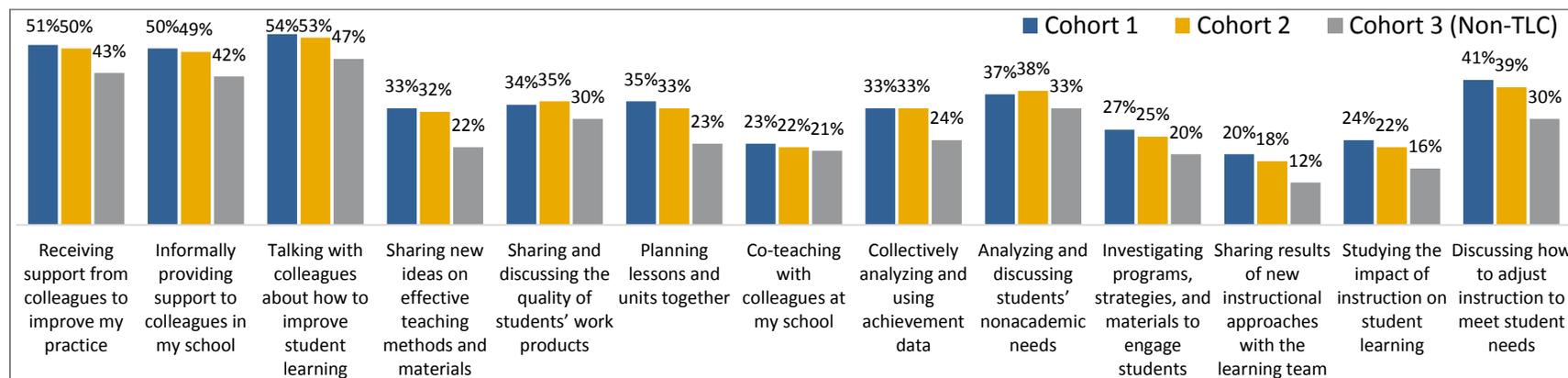
Among focus group participants, nearly all (97%) of the teacher leaders *agreed* or *agreed strongly* that TLC supported effective instruction. The focus groups indicated the program supports instruction through instructional coaches and more collaboration among teachers. In interviews, 6 out of 10 district administrators reported that the TLC program had improved teacher instruction in their school as a result of the use of instructional coaches. However, only 3 out of 10 administrators believed instruction improved solely because of more teacher collaboration. Interviewed

AEA staff believed TLC improved instruction by improving the instructional strategies used in the classroom (4 out of 12 AEA staff) and by embedding instructional coaches (4 out of 12 AEA staff).

Changes to the professional climate between teachers and administrators. TLC teachers and administrators indicated that TLC is positively affecting their professional structure. When asked a

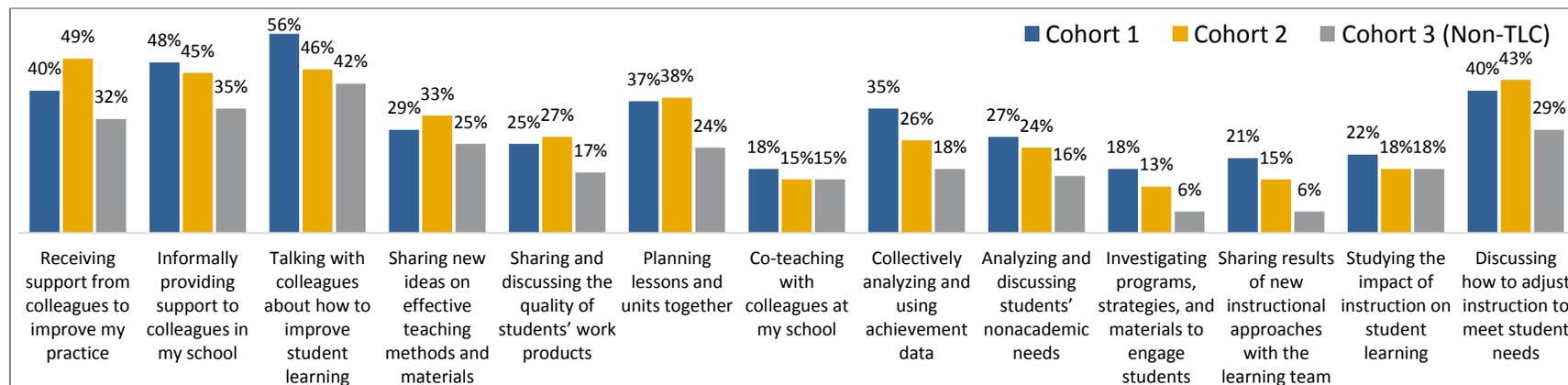
series of questions about whether the perceived TLC-related changes in professional climate were positive (e.g., teachers are better compensated, teachers have more opportunities and feel listened to, and time for collaboration has increased), scores for 72% of teachers and 90% of administrators fell in either the *agree somewhat* or the *agree strongly* categories. Perceived positive changes in professional climate were similar for teachers and administrators in Cohorts 1 and 2, although larger percentages of administrators, compared with teachers, indicated that the changes are positive.

Figure 5. Percentage of Surveyed Teachers Who Reported That Teachers Participate in the Following Collaboration Activities at Their School at Least Once a Week



Note: Sample sizes for surveyed teachers: $n_{c1} = 3,379$, $n_{c2} = 3,115$, $n_{c3} = 3,637$. Responses for *observing colleagues teaching practice* and *having colleagues observe my teaching practice* were endorsed by less than 10% of respondents, on average, and were not included in this figure.

Figure 6. Percentage of Surveyed Administrators Who Reported That Teachers Participate in the Following Collaboration Activities in Their School at Least Once a Week



Note: Sample sizes for surveyed administrators: $n_{c1} = 276$, $n_{c2} = 286$, $n_{c3} = 448$. Responses for *observing colleagues teaching practice* and *having colleagues observe my teaching practice* were endorsed by less than 10% of respondents, on average, and were not included in this figure.

Evaluation of Impacts on Student Achievement

To estimate the impact of TLC on student achievement in the first year of program implementation, we used a comparative interrupted time series (CITS) design.⁵ This design examined the historical performance of students in TLC districts (Cohort 1) and students in districts that had not yet implemented the program as of the 2014–15 school year (Cohorts 2 and 3) to predict post-TLC implementation student achievement

The CITS design uses the historical performance of students in TLC districts and students in districts that had not yet implemented the program to predict post-TLC implementation outcomes. The design does not require that comparison districts be identical to TLC districts.

outcomes. The initial analyses used student assessment scores in reading and mathematics from 2005–06 to 2014–15, standardized so that scores from different assessments were on the same scale.⁶

The design does not require that comparison districts be identical to TLC districts, and assumes that events that coincide with implementing TLC (e.g., policies implemented, assessment changes) equally affect TLC and non-TLC comparison districts. The CITS design is especially useful for evaluating enacted programs or policies in a statewide rollout over time.

Interim Findings on TLC Program Impact

This section describes interim findings about the impact of TLC on student achievement, including overall impact estimates based on data from nearly all students in Iowa followed by findings related to subgroups. We present our results in effect sizes, which convey the direction and magnitude of a relationship.⁷ Positive effect sizes indicate that the TLC districts outperformed non-TLC districts, while negative effect sizes indicate that non-TLC districts outperformed TLC districts. We also provide context on how observed effect sizes translate into score differences on the Iowa Assessments.

TLC districts improved slightly less on the Iowa Assessments in reading and mathematics in the first year of implementation for Cohort 1 compared to non-TLC districts.

Student achievement in the TLC Cohort 1 districts improved slightly less relative to their pre-intervention average, by about 0.04 standard deviations in reading and by about 0.05 standard deviations in mathematics, compared to pre–post difference for students in non-TLC districts in the first year of implementation (see Figure 7). Historically, the TLC Cohort 1 districts performed marginally below the state averages in reading and mathematics, while non-TLC districts performed marginally above. The difference between these groups widens in the 2014–15 year, with non-TLC districts improving slightly more than TLC districts. This difference is small, but statistically

⁵ CITS designs are among the strongest quasi-experimental designs for causal inference (Shadish, Cook, & Campbell, 2002; St. Clair, Cook, & Hallberg, 2014), and recent methodological studies have demonstrated their comparability to randomized controlled trials—the gold standard for measuring an intervention’s effect—using within-study comparisons (e.g., Hallberg, Williams, & Swanlund, 2015; Jacob, Somers, Zhu, & Bloom, 2016).

⁶ AIR standardized the assessment scores by grade, subject, and year. Since Iowa has multiple testing windows (fall, midyear, and spring) for data provided to AIR, the DE calibrated student scale scores to match the spring testing window.

⁷ Effect sizes here are presented in standard deviation units. The standard deviation is a measure of variation or dispersion around the mean, with larger values indicating greater variation and smaller values closer to zero indicating less.

significant.⁸ The difference translates into approximately 1–2 points on Iowa assessments that span about 200 points.⁹

Figure 7. Average Standardized Reading and Mathematics Scale Scores by School Year, TLC Cohort 1 and Non-TLC Districts



Note. Figure 1 presents the average standardized scores for reading and mathematics for TLC Cohort 1 districts and non-TLC districts (Cohorts 2 and 3). The overall mean in each year of data is zero. Values above zero indicate that group's average was above the state average. Values below zero indicate that group's average was below the state average. The models used to estimate TLC's impact on student achievement adjusted for student demographics, grade, and AEA.

⁸ With 10 years of Iowa's historical student achievement data for a large statewide sample, the analytic approach can detect such minor changes in academic performance.

⁹ We calculated the standard deviation for the raw Iowa Assessments scale scores in reading and in mathematics, pooling across grades and years, and then multiplied impact estimates by the scale score standard deviation values for an approximation of change in terms of scale score points on the Iowa Assessments.

Differences in achievement between TLC Cohort 1 districts and non-TLC districts varied slightly by district size.

We examined the impact of TLC on student achievement in Cohort 1 within six district size tiers and compare those impact estimates to other districts not in that size tier.¹⁰ Differences indicate that the effects of TLC vary by district size. In districts with 2,500 to 8,999 students, the TLC impact on mathematics was smaller compared to the other districts. Conversely, the TLC impact on mathematics was larger in districts with 1,000 to 2,499 students and in districts with 600 to 999 students compared to the other districts. Additionally, districts with 1,000 to 2,499 students saw a larger TLC impact on reading compared to the other districts. These differences were statistically significant but small and translated into about 1–3 points on the Iowa assessments. No significant differences were found for districts with 9,000 or more students, districts with between 300 and 599 students, or districts with 300 or fewer students enrolled, compared to other-sized districts.

Achievement gaps were mostly unchanged for student subgroups between TLC Cohort 1 districts and non-TLC districts.

In a similar fashion, we examined variation in TLC’s impact across different student subgroups, including English language learners, students eligible for free or reduced-price lunch, and students with an individualized education program (IEP). The achievement gaps in TLC Cohort 1 versus non-TLC districts remained unchanged for most student subgroups, meaning the effects of TLC were comparable for students in these subgroups compared to students who were not.

One small, significant difference was observed. The difference between TLC and non-TLC districts among students with an IEP was more positive than the difference for students without an IEP, by about 0.04 standard deviations, which represents a small, statistically significant decrease in the achievement gap between these two groups. This effect is also equal to approximately 1–2 points on the Iowa Assessments.

¹⁰ District size tiers examined: 9,000 or more students, 2,500 students to 8,999 students, 1,000 students to 2,499 students, 600 students to 999 students, 300 students to 599 students, and less than 300 students.

Conclusions and Implications

Teacher and administrator input on TLC implementation is consistent with expected progress in implementing the program's services, based on analyses conducted to date, though this may be due to preexisting conditions rather than influence of TLC. Compared to non-TLC districts, respondents from TLC districts were more likely to perceive greater availability, frequency, or quality of key aspects of the TLC program, including teacher leadership roles, professional development supports for teachers, and teacher collaboration. Among districts that had implemented TLC, respondents in Cohort 1 districts where implementation had occurred for a longer period of time often had more positive responses on average than respondents in Cohort 2 districts. Most teachers and administrators in TLC districts also indicated that the program was effective for improving instruction, teacher satisfaction, and professional climate.

The initial evaluation activities identified potential areas for improvement in program implementation. Some teachers in TLC districts were still not adequately familiar with the roles and supports provided by the program, and administrators tended to be more familiar than teachers were. No differences were observed between TLC and non-TLC districts in perceived opportunities for teacher leadership roles. And only small differences were found between TLC and non-TLC districts in the percentage of teachers who reported having teacher leader roles. As the program implementation continues, clearly defining and communicating about the roles and supports available for teachers at the local level may help expand teachers' understanding of the professional opportunities TLC is intended to provide.

Analyses of student achievement were based on data from the initial year of implementation for Cohort 1, and it is therefore too early to reliably discern the overall impact of TLC. Potential benefits of the TLC-related supports that districts put in place may not yet be apparent. The primary finding that student achievement in the TLC districts improved slightly less compared to achievement in non-TLC districts may not be meaningful in practice. As the achievement data become available for subsequent academic years, we will have the opportunity to look at second-year impacts for Cohort 1 and first-year impacts for Cohort 2.

Next Steps

This brief highlighted initial findings. AIR will complete analysis of data collected thus far (including data from surveys, interviews, and focus groups) and summarize the interim implementation findings in reporting to the DE in early 2017. For future evaluation activities, we will gather new data on program implementation in the 2016–17 academic year and examine effects of the TLC program on student achievement, using data from the 2015–16 and 2016–17 academic years.

References

- Battaglia, M. P., Hoaglin, D. C., & Frankel, M. R. (2009). Practical considerations in raking survey data. *Survey Practice*, 2(5). Retrieved from <http://www.surveypractice.org/index.php/SurveyPractice/article/view/176/html>
- Hallberg, K., Williams, R. T., & Swanlund, A. P. (2015). *Assessing the validity of comparative interrupted time series designs in practice: Lessons learned from two within-study comparisons*. Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, Miami, FL.
- Jacob, R., Somers, M.-A., Zhu, P., & Bloom, H. (2016). The validity of the comparative interrupted time series design for evaluating the effect of school-level interventions. *Evaluation Review*, 0193841X16663414. <http://doi.org/10.1177/0193841X16663414>
- Linacre, J. M. (2015). Winsteps® (Version 3.91.0) [Computer software]. Beaverton, OR: Winsteps.com.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference (2nd Ed.)*. Boston, MA: Houghton Mifflin.
- St. Clair, T. S., Cook, T. D., & Hallberg, K. (2014). Examining the internal validity and statistical precision of the comparative interrupted time series design by comparison with a randomized experiment. *American Journal of Evaluation*, 35(3), 311–327.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: Mesa Press.