



Research Memorandum

ETS RM–16-12

Understanding Consequential Assessment Systems for Teachers: Year 1 Report to the Los Angeles Unified School District

Courtney Bell

Leslie Stickler

Nathan Jones

Shuangshuang Liu

Jennifer Lewis

Monica McLeod

Yi Qi

November 2016

ETS Research Memorandum Series

EIGNOR EXECUTIVE EDITOR

James Carlson
Principal Psychometrician

ASSOCIATE EDITORS

Beata Beigman Klebanov
Senior Research Scientist

Heather Buzick
Research Scientist

Brent Bridgeman
Distinguished Presidential Appointee

Keelan Evanini
Research Director

Marna Golub-Smith
Principal Psychometrician

Shelby Haberman
Distinguished Presidential Appointee

Anastassia Loukina
Research Scientist

Donald Powers
Managing Principal Research Scientist

Gautam Puhan
Principal Psychometrician

John Sabatini
Managing Principal Research Scientist

Matthias von Davier
Senior Research Director

Rebecca Zwick
Distinguished Presidential Appointee

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Memorandum series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Memorandum series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

**Understanding Consequential Assessment Systems for Teachers:
Year 1 Report to the Los Angeles Unified School District**

Courtney Bell, Nathan Jones, Jennifer Lewis, Yi Qi, Leslie Stickler,
Shuangshuang Liu, and Monica McLeod

Educational Testing Service, Princeton, New Jersey

November 2016

Corresponding author: C. Bell, E-mail: cbell@ets.org

Suggested citation: Bell, C., Jones, N., Lewis, J., Qi, Y., Stickler, L., Liu, S., & McLeod, M. (2016). *Understanding consequential assessment systems for teachers: Year 1 report to the Los Angeles Unified School District* (Research Memorandum No. RM-16-12). Princeton, NJ: Educational Testing Service.

Find other ETS-published reports by searching the ETS ReSEARCHER
database at <http://search.ets.org/researcher/>

To obtain a copy of an ETS research report, please visit
<http://www.ets.org/research/contact.html>

Action Editor: James Carlson

Reviewers: Joseph Ciofalo and Jonathan Steinberg

Copyright © 2016 by Educational Testing Service. All rights reserved.

ETS, the ETS logo, and MEASURING THE POWER OF LEARNING. are registered trademarks of Educational Testing Service (ETS). All other trademarks are the property of their respective owners.



Abstract

This report summarizes what was learned from the first year of the Understanding Consequential Assessment Systems of Teaching (UCAST) study. The study seeks to understand how administrators learn to use the observation portion of the Los Angeles Unified School District's (LAUSD) consequential teacher evaluation system for teachers. The report describes the 2012–2013 implementation in which more than 1,000 LAUSD building administrators were trained and certified to implement the observations that are a part of the Teacher Growth and Development Cycle (TGDC). Each administrator worked with one teacher during this gradual implementation year. Using mixed methods, the report summarizes administrators' background characteristics, perceptions of training and of the TGDC system, results of the training and certification, and use of the TGDC system. Recommendations for revisions to the system are also included.

Key words: Teacher evaluation, observations, teaching quality, administrators, teacher assessment, accountability, policy, Teacher Growth and Development Cycle (TGDC)

Acknowledgments

We would like to thank the administrators, trainers, central office staff, and focus administrators who generously allowed us access to their learning in the 2012–2013 school year. Their commitment to providing honest appraisals, helpful feedback, and support has made this work possible.

We would also like to thank the core team of LAUSD staff and contractors who regularly responded to our e-mails and answered their phones when we had questions. We would especially like to thank Donna Muncey, Rachel Bonkovsky, Brian Lucas, Drew Furedi, Lauren Gregor, Duffy Miller, John Armbrust, Maria Salgado, Seth Cutler, Karissa Yee, and Jolene Chavira. Your commitment to learning from this work is commendable.

Colleagues at Educational Testing Service were invaluable in the preparation of this report and the administration of the grant. Debi Ackerman, Jennifer Caterson, Evelyn Fisch, Darlene Rosero, and Heather Walters were helpful, kind, and ready to deal with any problem we devised.

Finally, a special thank you to the W. T. Grant Foundation for its generous support of this study.

Table of Contents

	Page
List of Tables	vi
Los Angeles Unified School District’s Teacher Growth and Development Cycle, 2012–2013	2
Who Participated in 2012–2013 Training?	3
Participation Summary.....	3
Description of Participants.....	3
Who Are the Focus Administrators?.....	4
Summary of Findings.....	4
Recruitment of Focus Administrators	4
Focus Administrator Characteristics	5
How Did the 2012–2013 Trainings Go?	6
Summary of Findings.....	6
Training Description	6
Administrators’ Views of Training	10
Summer and School Year Trainings in 2012–2013	13
Training Changes for 2013–2014	13
How Did Administrators Perform on the Certification Assessment?	14
Summary of Findings.....	14
What Aspects of the Teaching and Learning Framework Are Most Challenging?	15
Overall Certification Performance	15
What Predicted Certification and Training Satisfaction?	19
Summary of Findings.....	19
Background	20
Predicting Administrator Satisfaction With Training	20
Predicting Certification Outcomes.....	22
How Did the 2012–2013 Observations With Teachers Go?.....	25
Summary of Findings.....	25
Background	25
Finding a Focus Teacher	25
Experiences With Formal and Informal Observations.....	27

Reflections on Future Training Needs	30
How Did Administrators' Views Change Over the School Year?	31
Summary of Findings.....	31
Administrators' Support of Teacher Growth and Development Cycle Human Capital	
Goals	32
Administrators' Views of the Teaching and Learning Framework	34
Administrators' Job Efficacy	35
How Did Administrators Use the Teacher Growth and Development Cycle in the 2012–2013	
School Year?	36
Summary of Findings.....	36
Background	36
Interactions With the Teacher Growth and Development Cycle	36
Number and Types of Observations	37
Supports Used to Complete Observations	37
Use of the Teaching and Learning Framework During Observations	39
Score Creation Process	39
Evaluating a Teacher.....	42
Summary Findings and Recommendations	45
Training.....	45
Certification	46
Conducting Observations and Using the Teacher Growth and Development Cycle	47
Changes in Administrators' Views	49
Recommendations Summary	50
Training.....	50
Certification	50
Conducting Observations and Using the Teacher Growth and Development Cycle	51
How Was This Report Prepared?.....	52
Training Surveys for All Administrators	52
Surveys, Interviews, and Other Data for Focus Administrators	53
Other Data Sources	56
Data Analysis	57

Writing	58
Glossary	59
Evaluation Context.....	59
Personnel.....	60
References.....	61
Appendix A. Descriptive Results.....	62
Descriptive Tables for Posttraining Survey	73
Descriptive Tables for End-of-Year Survey	81
Appendix B. Analyses	86
Appendix C. Year 1 Study Instruments	91
Expectations for Training and Prior Experiences	93
Focus Administrator Spring Interview Protocol	118
Training.....	119
Perception of the Evaluation Policy.....	119
Los Angeles Unified School District Administrator End-of-Year Survey	121

List of Tables

	Page
Table 1. Occupational Roles of Administrators Who Returned Presurveys.....	4
Table 2. Teaching Certification Areas of Administrators Who Returned the Pretraining Survey .	4
Table 3. Administrators' Views of Training.....	11
Table 4. Miller and Infelice Certification Rates (Overall and by Component)	16
Table 5. Certification Rates by Quantitative Measures (Existing Composite and Accuracy Categories)	17
Table 6. Pretraining Survey Measures Used to Predict Training Satisfaction.....	21
Table 7. Regression Coefficients for Predicting Training Satisfaction Using Pretraining Measures of Beliefs and Expectations (Standard Errors).....	23
Table 8. Regression Coefficients Predicting Certification Using Pre- and Posttraining Measures of Beliefs and Expectations (Standard Errors).....	24
Table 9. Administrators' Experience With Focus Teachers in 2012–2013, End-of-Year Survey Sample.....	26
Table 10. Statistics on Formal and Informal Observations Administrators Did With Focus Teachers in 2012–2013, End-of-Year Survey Sample	27
Table 11. Focus Administrator Agreement Regarding Informal Observation, Mid-Year Survey Sample.....	28
Table 12. Administrators' Views of Training Outcomes, End of-Year Survey Sample	30
Table 13. Focus Administrators' Rankings of Professional Learning Needs	31
Table 14. Administrators' Mean Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes Before and After Training	33
Table 15. Administrators' Mean Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes on Posttraining and End-of-Year Surveys	33
Table 16. Administrators' Mean Views of the Teaching and Learning Framework on Posttraining and End-of-Year Surveys.....	34
Table 17. Administrators' Mean Views of Training Outcomes on Posttraining and End-of-Year Surveys.....	34
Table 18. Administrators' Mean Views of Their Job Efficacy and Focus on Pretraining and End- of-Year Surveys	35

Table 19. Extent of Resource Use to Support Teacher Growth and Development Cycle Implementation, 2012–2013	38
Table 20. Survey Schedule and Response Rates	53
Table A1. Occupational Roles of Administrators, Pretraining Survey Sample.....	62
Table A2. Number of Years Administrators Served in Same Job Prior to This School Year, Pretraining Survey Sample	62
Table A3. Grade Levels Administrators Work With Regularly, Pretraining Survey Sample	62
Table A4. Certification Areas of Administrators Who Returned Presurveys.....	63
Table A5. Certification Areas of Administrators, Pretraining Survey Sample.....	63
Table A6a. Focus Administrators' Weekly Time Allocation by Area	63
Table A6b. Focus Administrators: Important Aspects of Their Job, Other Than Listed.....	64
Table A6c. Focus Administrators: Time-Consuming Aspects of Their Job, Other Than Listed .	64
Table A7. Administrators' Expectations for Learning in 2012–2013 Prior to Training, Pretraining Survey Sample	65
Table A8. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data Sources for Improving Instruction, Pretraining Survey Sample	65
Table A9. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes, Pretraining Survey Sample.....	66
Table A10. Frequency of Specific Evaluation Practices Prior to 2012–2013, Pretraining Survey Sample.....	66
Table A11. Focus Administrator Agreement Regarding Observation Practices Prior to Teacher Growth and Development Cycle Training.....	66
Table A12. Frequency of Specific Evaluation Practices Prior to Implementing the Teacher Growth and Development Cycle.....	67
Table A13a. Administrators' Agreement With Statements About What Makes for Effective Teaching, Pretraining Survey Sample.....	68
Table A13b. Administrators' Agreement With Statements About What Makes for Effective Teaching, Summer Versus School Year Training Participants	69
Table A14. Administrators' Evaluation of Full-Time K–12 Classroom Teachers in 2011–2012, Pretraining Survey Sample	70
Table A15a. Administrators' Perceptions of Job Efficacy, Pretraining Survey Sample	70

Table A15b. Administrators' Perceptions of Job Efficacy, Summer Versus School Year Training Participants.....	71
Table A16. Administrators' Note-Taking Approach, Pretraining Survey Sample.....	71
Table A17. Administrators' Responses to Question 1 on the Mid-Training Survey.....	71
Table A18. Administrators' Responses to Question 2 on the Mid-Training Survey.....	72
Table A19. Administrators' Responses to Question 3 on the Mid-Training Survey.....	72
Table A20. Administrators' Responses to Question 4 on the Mid-Training Survey.....	72
Table A21. Administrators' Responses to Question 5 on the Mid-Training Survey.....	72
Table A22. Consistency Between the Training and Administrators' Expectations, Posttraining Survey Sample	73
Table A23. Administrators' Beliefs About Applying Training to Their Own Schools, Posttraining Survey Sample.....	73
Table A24. Administrators' Evaluation of Full-Time K–12 Classroom Teachers in 2012–2013, Posttraining Survey Sample.....	74
Table A25. Administrators' Views of Training Outcomes, Posttraining Survey Sample	74
Table A26a. Administrators' Views of the Teaching and Learning Framework, Posttraining Survey Sample	75
Table A26b. Administrators' Views of the Teaching and Learning Framework, Summer Versus School Year Training Participants	76
Table A26c. Administrators' Views of the Teaching and Learning Framework, by Job Role	77
Table A26d. Administrators' Views of the Teaching and Learning Framework, by Instructional Level	79
Table A27. Administrators' Views of the Utility of Teacher Growth and Development Cycle Observation Component for Various Purposes, Posttraining Survey Sample.....	80
Table A28. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes, Posttraining Survey Sample.....	80
Table A29. Administrators' Views of Training Outcomes, End-of-Year Survey Sample	81
Table A30. Number of Formal and Informal Observations Administrators Have Done With Focus Teachers in 2012–2013, End-of-Year Survey Sample.....	81
Table A31. Total Number of Hours to Complete a Formal Observation Cycle, End-of-Year Survey Sample	81

Table A32. Administrators' Impression of the Teaching and Learning Framework, End-of-Year Survey Sample	82
Table A33. Administrators' Views of the Utility of Teacher Growth and Development Cycle Observation Component for Various Purposes, End-of-Year Survey Sample	82
Table A34. Administrators' Perceptions of Job Efficacy, End-of-Year Survey Sample.....	83
Table A35. Administrators' Implementation of Teacher Growth and Development Cycle (TGDC) Training, End-of-Year Survey Sample.....	83
Table A36. Challenges Associated With Observing and Evaluating Specific Teacher Groups...	83
Table A37. Focus Administrators: Comments on Teacher Growth and Development Cycle (TGDC)—Open-Ended Responses.....	84
Table A38a. Certification Results (by Occupational Role)	85
Table A38b. Certification Results (by Grade Level)	85
Table B1. Pretraining Survey Measures Used to Predict Training Satisfaction	86
Table B2. Regression Coefficients Predicting Training Satisfaction Using Pretraining Measures of Beliefs and Expectations	86
Table B3. Predicting Certification Using Pre- and Posttraining Measures of Beliefs and Expectations.....	87
Table B4. Pearson Correlation Among Certification Measures	87
Table B5. Element-Level Reliability Statistics (Full Sample).....	88
Table B6. Element-Level Reliability Statistics (by Job Role).....	89
Table B7. Element-Level Reliability Statistics (by Instructional Level).....	90

In this report we summarize the first year of data collection in the Understanding Consequential Assessment Systems for Teachers (UCAST) study, funded by the W. T. Grant Foundation. The goal of UCAST is to understand the process by which local administrators in the Los Angeles Unified School District (LAUSD) learn to conduct Teacher Growth and Development Cycle (TGDC) observations.

We present findings from our data collection across the 2012–2013 school year, beginning with the initial training process for administrators, their performance on and perceptions of the observation certification assessment, their experiences conducting and using the observation system across the school year, and the relationship between the TGDC and other district initiatives. We have drawn on various sources of data in this report, including (a) pre-, mid-, and posttraining surveys of all participating administrators; (b) TGDC training field notes recorded by the research team; (c) observation certification assessment results; (d) interviews conducted with a group of 42 focus administrators during training and at the end of the school year; (e) additional focus administrator surveys; (f) think-aloud exercises completed by the 42 focus administrators; (g) focus groups with principals and instructional directors; and (h) interviews with central office staff.

The following research questions are the focus of this report:

1. Who participated in 2012–2013 training?
2. Who are the focus administrators?
3. How did the 2012–2013 trainings go?
4. How did administrators perform on certification?
5. What predicted training satisfaction and attainment of certification?
6. How did the 2012–2013 observations with teachers go?
7. How did administrators' views change over the school year?
8. How did administrators use the TGDC in the 2012–2013 school year?

We describe how we investigated these questions then present findings and implications. We conclude the report with a series of recommendations for LAUSD as they begin planning their implementation of TGDC in 2013–2014. Attached to the end of the report is a section

documenting our methodology as well as appendices with our study instruments and supplemental analyses.

Los Angeles Unified School District's Teacher Growth and Development Cycle, 2012–2013

LAUSD is the second largest public school district in the United States, with more than 800 schools and approximately 670,000 enrolled students. The student population is racially and ethnically diverse; teachers are similarly diverse. More than 76% of students are eligible for free or reduced-price lunches. Increasingly, LAUSD has occupied a visible position in national conversations about teacher evaluation, with its efforts to train administrators as classroom observers being closely watched by other districts and states. The size and diversity of LAUSD allowed us to explore variation in how administrators think about and perform teacher evaluations.

In 2012–2013, LAUSD implemented a “practice year” of its new teacher evaluation system, the TGDC, which was initially rolled out as a pilot with a limited number of administrator and teacher volunteers during the 2011–2012 school year. The TGDC had previously been referred to as the Educator Growth and Development Cycle, but with the development of the School Leader Growth and Development Cycle for administrators and an evaluation process for central office officials, the Educator Growth and Development Cycle came to represent the larger evaluation framework for various professionals. LAUSD partnered with Teaching Learning Solutions (TLS) to provide training to administrators. TLS also developed the online system through which administrators recorded ratings and observational notes for each TGDC evaluation.

The observation protocol for the TGDC, called the Teaching and Learning Framework (TLF), is a modified version of Danielson's Framework for Teaching (Danielson & McGreal, 2000). Danielson's instrument is reported to be the most widely used observation protocol in the country. The new instrument, the TLF, has been aligned to the California teaching standards and modified to reflect the values of LAUSD stakeholders.

The TGDC is a significant departure from past teacher evaluation practices. It establishes a vision of high-quality instruction that may or may not be shared by teachers, administrators, and other education professionals in the district. TGDC requires administrators to be trained and become certified in ways that are unprecedented in LAUSD. Quality control mechanisms and the

early and ongoing evaluation of TGDC are commendable given how little researchers and policy makers know about reforming human capital systems in education.

Who Participated in 2012–2013 Training?

Participation Summary

A total of 998 administrators, mostly principals and assistant principals, received training on the TGDC during the 2012–2013 school year. Of those to whom we sent pretraining (820 administrators) and posttraining (966 administrators) surveys about their training experience and opinions about TGDC, most responded (677 administrators, or 82.6%, responded to the pretraining survey; 633 administrators, or 65.5%, responded to the posttraining survey).

Description of Participants

To better understand administrators' roles, characteristics, and beliefs about teaching and learning, we sent pre-, mid-, and posttraining surveys to all administrators who registered for training ahead of time. Not all administrators registered in advance or attended the training session assigned, however, which reduced the survey sample to 820 administrators for the pretraining survey and 966 administrators for the posttraining survey. Recruitment methods, survey administration, and response rates are described more fully in the section titled *How Was This Report Prepared?*

Of the administrators who completed surveys, the majority (88% total) were either principals (64%) or assistant principals (24%), as shown in Table 1. Thirty-one administrators, or about 5% of administrators who completed the pretraining survey, were central office staff. Central office staff included individuals with job titles such as specialized director or least restrictive environment specialist. A small proportion of administrators were instructional directors (1%).

On average, administrators had 6 years of experience. The majority of administrators also held teaching certification in a wide range of areas (see Table 2). Most were certified as elementary educators. There were fewer administrators certified in mathematics or science. More details about administrators' certifications are presented in Table A4.

Who Are the Focus Administrators?

Summary of Findings

We recruited 42 focus administrators to provide more detailed information about the training and implementation experiences of the TGDC. Focus administrators completed additional surveys, interviews, and think-aloud exercises to help us delve into their experiences learning to use the TGDC across the practice year. We have complete data across all instruments from 38 of the 42 focus administrators. Overall, the focus administrator sample is similar to other administrators who participated in the study, with some small differences.

Table 1. Occupational Roles of Administrators Who Returned Presurveys

Occupational role	Overall	Summer training	School year training
Principal	432 (63.8%)	399 (75.4%)	33 (22.4%)
Assistant principal	164 (24.2%)	110 (20.8%)	54 (36.7%)
Instructional director	9 (1.3%)	8 (1.5%)	1 (0.7%)
Central office staff	31 (4.6%)	4 (0.8%)	27 (18.4%)
Other	40 (6.0%)	8 (1.5%)	32 (21.8%)
Missing	1 (0.1%)		
Total	677 (100.0%)	529 (100.0%)	147(100.0%)

Note. $N = 677$.

Table 2. Teaching Certification Areas of Administrators Who Returned the Pretraining Survey

Teaching certification area	Percentage	Count
Elementary education	63.5	429
Mathematics	13.8	93
English/language arts	20.7	140
Science	10.4	70
Social studies	17.0	115
Total		847

Note. $N = 676$. Percentages do not sum to 100, and the total count does not match N because some administrators hold multiple certifications.

Recruitment of Focus Administrators

To better understand TGDC training and initial implementation, we focused more intensive data collection on a subgroup of 42 administrators. These focus administrators completed an extensive online questionnaire after training and three short bimonthly surveys throughout the school year. They also participated in several one-on-one interviews and completed think-aloud exercises. Focus administrators were recruited during the first training session in Summer 2012, which took place during the week of July 9, 2012. Research team members visited four training sites to introduce the evaluation study and recruit participants on-

site. To compensate focus administrators for their additional effort, they received \$500 for their participation in 2012–2013, paid with grant monies. Across three training sites, 121 administrators volunteered to participate as focus administrators. Each training site had two or three classrooms of participants, but we selected one classroom at each site to serve as a focal classroom. Across the three classrooms, 42 focus administrators were selected to participate in our study. Focus administrators were selected through stratified random sampling, stratifying by job type (principal, instructional director, etc.), and training site.

Focus Administrator Characteristics

Focus administrators were similar to the larger population of administrators who returned surveys; however, they differed in being slightly less experienced, including more people certified to teach English/language arts and science but fewer people certified with an elementary certificate. There were also slightly more focus administrators working in secondary settings as compared with all administrators.

More than half of the focus administrators were principals, with another quarter serving as assistant principals, four serving as instructional directors, and three serving as instructional coaches or instructional specialists. They had an average of 4.4 years in their current professional roles, including time in LAUSD and any other district. Assistant principals had spent the greatest amount of time in their current roles, averaging 7.9 years. Approximately 40% of the focus administrators worked in elementary schools, 20% of the focus administrators worked in middle schools, and 33% held positions in high schools. The four instructional directors worked across grade levels. Slightly more than half of the focus administrators were female.

Thirty-two focus administrators indicated that they were certified to teach, with 69% certified in teaching English learners, which was disproportionately higher than the larger sample of administrators who returned surveys. Half of focus administrators were certified in elementary education. One third of the focus administrators were certified in English/language arts, 25% held certificates for social studies, 25% held science credentials, and 19% were certified in mathematics.

Focus administrators had similar expectations of the training as the larger group of administrators but were more eager to learn how observations fit into the TGDC and how to help teachers improve (e.g., by matching teachers to appropriate professional development based on their observation performance or helping resistant teachers learn to improve their practice).

Compared to the larger group, a greater percentage of focus administrators reported that the TGDC data would be useful for teacher and school improvement. A higher percentage of focus administrators than administrators overall were optimistic that they could focus TGDC on the quality of instruction and learning in each classroom they observed.

How Did the 2012–2013 Trainings Go?

Summary of Findings

Overall, administrators felt that they learned how to use the TLF and how observations fit into the TGDC. Broadly, administrators felt that they learned skills directly related to the use of the evaluation tool, such as taking accurate notes and assigning scores, but learned less about implementing observations in practice, such as strategies for working with teachers who need more professional development and/or those who may be resistant to improvement. There was evidence that the content of training evolved over the school year to include a stronger focus on coaching, on distinguishing between levels of performance, and on the language of the TLF.

Training Description

Background. Initial TGDC trainings were conducted by staff from TLS over a 5-day period. Four days of instruction were devoted to (a) the TLF, (b) the technology platform for TGDC observations, and (c) broader facets of teacher evaluation in LAUSD. Administrators watched four to five videos of classroom practice during training and rated them on elements across each of the five TLF standards; they practiced scripting teachers' and students' behaviors during observations and learned to rate increasing numbers of standards over the week. Training also included discussion of what each standard meant.

UCAST research team members attended five trainings throughout 2012–2013—three in the summer and two during the school year. The content covered during summer training and school-year trainings differed slightly but was largely similar. Summer trainings began at 8:00 A.M. and lasted until 3:00 P.M. each day, with a 1-hour break for lunch and a 15-minute break in the morning. The last day of the training was dedicated to certification activities.

Training, Day 1. Training sessions began with conversations about participants' hopes, fears, and current understandings of the TGDC. After this conversation, administrators watched a video from the district superintendent explaining the role of evaluation in improving performance in LAUSD schools. Trainers then provided an overview of the TGDC process and

of the TLF, showing administrators how to navigate their training manuals. Additional topics included the context for adopting TGDC, research background for the TLF, and alignment of the Common Core State Standards (CCSS) with the TLF.

Trainers emphasized the role of evidence in conducting TGDC evaluations by providing criteria for evidence to support TLF ratings and by explaining the difference between evidence and opinion. This included a discussion of the kinds of evidence that can be gleaned from an observation, the objectiveness of evidence, and what types of evidence can justify lesson ratings. When taking notes during an observation, administrators were urged to do “verbatim scripting.” Trainers encouraged administrators to “leave behind biases” and record “just the facts.”

In the afternoon of the first day, administrators watched a video of a ninth-grade algebra lesson and practiced rating the elements of Standard 3 (Standards-Based Learning Activities). Administrators then watched a video of a kindergarten lesson and practiced rating elements of Standard 2 (Designing Coherent Instruction) as well as Standard 3. The length of discussion and elaboration of these elements varied between training sites. Observations by the research team suggested that time constraints required trainers to move on once they determined that administrators held a sufficient shared understanding of the elements and of applying them.

Training, Day 2. Most of the second day emphasized learning to work in My Professional Growth System (MyPGS), a data platform provided by Truenorthlogic. MyPGS was designed to be the primary medium of communication about TGDC activities between administrators and teachers. TGDC ratings are recorded in this platform. MyPGS training focused on entering data and navigating the system. Administrators learned how to attach their scripted notes to an observation and to tag segments of the script, that is, categorize specific pieces of evidence by the corresponding element on the TLF.

The trainers described Standard 1 (Planning and Preparation) and each of its elements, providing relatively brief answers to practical questions about conducting TGDC observations. Administrators had the opportunity to explore Standard 1 in more depth via group activities to generate effective or highly effective examples of classroom practice for each focal element. Small groups reported their examples to the whole group of administrators.

Administrators reviewed and entered parts of a sample lesson plan into the MyPGS system. While they did so, they also participated in a discussion about preobservation meetings

with teachers that focused on supporting the growth and development of teachers and the importance of presuming positive intent on teachers' parts.

In the afternoon of Day 2, participants viewed a video of a second-grade mathematics lesson and rated the lesson on TLF Standards 2 and 3. After completing their ratings, administrators discussed the quality of their evidence: Was it objective? Was it representative of the lesson? Was the evidence well aligned with the TLF? Participants also talked about the teachers' questions of students and how these affected their ratings. Some training sessions completed the day by reading *Rethinking Teacher Evaluation in Chicago* (Sartain et al., 2011), while others ran out of time. Ratings were submitted via the MyPGS system for feedback to administrators about their efforts.

Training, Day 3. Day 3 began with a discussion of administrators' feedback on their ratings of the second-grade mathematics lesson they viewed the previous day. This conversation addressed the meanings of the elements under each standard, interpretations of teacher moves, accuracy of scripting while performing an observation, and the value of the evaluation system overall. Many administrators were concerned about scripting adequately, seeing the lesson versus taking verbatim notes, and finding sufficient evidence to assign ratings. Some training sessions provided time to read an excerpt from *Ensuring Accurate Feedback from Observations* (Jerald, 2012).

Administrators then practiced rating a full observation cycle for an eighth-grade English language arts lesson. They entered lesson plan information into MyPGS and then rated on Standards 2, 3, and 5a. This exercise included watching a video-recorded postobservation conference with the eighth-grade teacher, which initiated discussion about holding postobservation conferences. In some sessions, administrators role-played their own postlesson conferences with this teacher.

Training, Day 4. At some training sites, Day 4 opened with a conversation about developing trust and having productive professional conversations as an instructional coach. Some groups read Chapter 5 of *Results Coaching: The New Essential for School Leaders* (Kee, Anderson, Dearing, Harris, & Shuster, 2010).

Administrators moved on to tagging evidence from a video-recorded fourth-grade English language arts lesson. They continued to voice concerns about being able to type quickly enough to script the lesson and having time to look up to see what was happening in the

classroom. In one session, administrators commented on differences between instruction that is valued by the TLF and “normal” instruction for many teachers. After another brief discussion of postobservation conferences, discussion returned to the logistics and mechanics of rating. By lunchtime, administrators were ready to compare their ratings to those assigned by master raters.

In the afternoon, administrators talked more about Standard 1 (Planning and Preparation), unpacking the elements and their meanings. After this, they had the option of being coached through more practice ratings or through beginning certification.

Training, Day 5. Certification exercises were held on the fifth and final day of training. Administrators who had not watched the 45-minute, video-recorded lesson for the certification assessment on Day 4 saw it on Day 5; administrators were not allowed to view the certification video more than once. After watching the video, administrators could take up to 3 days to complete the certification assessment online from any location. Many opted to stay on-site and complete the certification assessment right away. Administrators rated the video-recorded lesson on Standards 1, 2, 3, and 5 via MyPGS.

Training emphasis. About 70% of time in training was focused on learning the mechanics of the MyPGS platform, how to record evidence, and how to align evidence with focus elements of the TLF. Administrators watched four to five videos total during training, practicing scripting teacher and student behaviors while observing and adding more standards as they progressed. Training sessions required administrators to elaborate on the meanings of the standards and their constituent elements and to think of examples they might see in classrooms. Sessions briefly touched on the ideas that undergird the TLF, including 21st-century skills, cognitive engagement, and constructivist learning. Researchers noted that all videos shown during the training were in mathematics and English language arts, despite administrators’ needs to rate all subjects during the school year.

Administrators raised many process-oriented questions about TGDC, for example, how many teachers will we be rating each year when this system is fully operational, what resources will be available when we identify specific, professional support needs among teachers, and what will happen if an administrator does not pass the certification assessment. These questions were addressed during training but were not fully resolved.

Given the research team’s collective experience training administrators to score, it struck us that there was very little discussion of why particular videos were rated at a certain level (e.g.,

developing, effective). In our experience, delving into these distinctions and justifications has helped raters understand both the observation system and approaches to rating in general. Making time to discuss these issues refines judgment so that ratings are reliable. Our observation that this type of discussion was missing from the training, however, should not be taken as a critique. Very little time was wasted during training, and administrators covered a great deal each day. On the other hand, at the conclusion of training, administrators still struggled to apply the TLF accurately and assign ratings similar to those of master raters—or other administrators. This may be a result of not making time to discuss why different videotaped lessons were rated at certain levels.

It is worth noting that there was variation across the trainings we observed. In addition to the style of particular facilitators, there were substantive differences in what was taught and how much emphasis different aspects of the training were given.

Administrators' Views of Training

In this section, we report administrators' views of the training based on survey responses, our field notes from the training sessions, and interviews with focal administrators. When possible, we triangulated data across sources. Descriptive results from survey instruments and certification data are summarized in Appendix A. Additional analysis results are presented in Appendix B.

An important goal of surveying administrators prior to their training was to examine their preexisting beliefs and practices regarding evaluation. This background knowledge was expected to shape what they learned from their training and how they applied their learning to implement the TGDC. We also explored administrators' existing beliefs and practices surrounding evaluation (Table A10). To summarize, administrators reported many experiences observing teachers informally and pointing out instructional strengths and weaknesses during postobservation conferences. They had less experience working with teachers to plan lessons, review lesson plans, or suggest resources to improve their instruction.

Administrators' beliefs about teaching quality also shaped what and how they learned. Table A13a summarizes administrators' perceptions about effective teaching. Many administrators ($N = 122$) wrote in additional skills not mentioned, suggesting that the questionnaire did not capture all of the work administrators felt that effective teachers do. The "other" skills and competencies they listed were varied and comprehensive. In general, skills

were instructional in nature and ranged from using higher levels of Bloom’s taxonomy to backward planning to using portfolios and cooperative groups.

A great deal of the TGDC system’s validity rests on what principals learned during training. The training conveyed a view of high-quality instruction, showed administrators how to rate classroom practice, and touched on the ideas on which the evaluation had been built, working with teachers to improve instruction, and balancing the new evaluation system with existing duties. Administrators’ perceptions of the training probably shaped their rating experiences moving forward. Administrator feedback about training experiences can be used to refine future trainings LAUSD may offer.

Table 3 summarizes administrators’ views of training, with responses ranging from 1 (*not at all*) to 4 (*a lot*). Our results indicated that most administrators (81%) believed they learned a lot about how observations fit into the TGDC. More than half (58%) learned a lot about using the MyPGS system for rating, and a similar portion (55%) said they learned how to rate accurately. Most administrators (81%) said they learned some or a lot about how to take good observation notes. Most administrators (84%) also learned some or a lot about giving feedback to teachers. Survey results on administrators’ views of the TLF by job role and grade level are reported in Tables B26c and B26d.

Table 3. Administrators’ Views of Training

View	Not at all	A little	Some	A lot
How observations fit into the TGDC	0.0% (0)	1.8% (11)	16.9% (106)	81.4% (511)
How to score observations more accurately	0.2% (1)	4.3% (27)	40.4% (253)	55.2% (346)
How to manage the TGDC with my other responsibilities	36.5% (229)	29.0% (182)	25.8% (162)	8.6% (54)
How to take good notes when observing	3.8% (24)	15.3% (96)	41.6% (261)	39.2% (246)
How to use MyPGS, the online teacher evaluation tool	2.2% (14)	8.1% (51)	31.9% (200)	57.7% (362)
How to give teachers feedback based on observed performance	1.6% (10)	14.1% (88)	41.2% (258)	43.1% (270)
How to match teachers to appropriate professional development based on their observed performance	21.4% (133)	25.8% (160)	35.4% (220)	17.4% (108)
How to help people who are resistant to change learn to improve their practice	25.5% (160)	30.9% (194)	34.1% (214)	9.6% (60)
How to effectively manage all the strengths and weaknesses of my staff	26.6% (167)	28.7% (180)	35.7% (224)	9.1% (57)

Note. Responses listed as percentages, with frequencies in parentheses. TGDC = Teacher Growth and Development Cycle.

Table 3 also highlights areas in which administrators reported learning less. Most of these areas were not a direct focus of the TGDC training. For example, almost half (47%) of administrators learned “little” or “not at all” about matching teachers with professional development opportunities appropriate to their needs. Even more (56%) reported learning little or not at all about helping teachers who are resistant to change, and a similar proportion (55%) learned little or nothing about effectively managing the strengths and weaknesses of their staff. Almost two thirds (66%) of administrators said they learned little or nothing about balancing TGDC with other job responsibilities. These topics were not covered in depth during the training, despite their importance for translating evaluation results into instructional improvements.

Focus administrators expressed concerns that echoed those found in the full sample. Specifically, when we analyzed focus administrator interviews about training, three similar themes emerged to support the survey responses. Specifically, the three themes administrators expressed in interviews were as follows:

- The new evaluation system is objective.
- The evaluation system will place many new demands on principals.
- The evaluation system will identify and promote good teaching.

Focus administrators reported confidence as well as experience conducting observations and supporting teacher growth (see Table A11). They expressed great comfort about applying the TLF across different grade levels and subject areas. One administrator anticipated that using the TGDC in practice would be easier due to familiarity with the students and the classroom context, as opposed to the training videos, which provided little context for the recorded lessons.

Like administrators in the full sample, the focus administrators were concerned about finding enough time for TGDC while juggling other essential tasks, and several were worried about technical aspects of the system, including verbatim scripting of lessons. One focus administrator wondered how much instruction would have to be modeled for teachers to help them learn how to teach in alignment with TLF. Focus administrators also noted that teachers differ and that, as administrators, they needed to focus their attention strategically to make the best use of limited time to conduct observations. Table A6a details how focus administrators reported spending their time.

Focus administrators varied in how they decided to allocate time spent on evaluation. Some decided to focus their time on teachers they believed most likely to be serving students poorly. Others tried to spread their time equally across all participating teachers. Despite varying approaches, focus administrators generally reported that no one group of teachers was substantially more challenging to evaluate (see Table A36). Some focus administrators did say it was more challenging to work with teachers with whom they had had conflict (31%), very experienced teachers (18%), or special education teachers (15%).

Summer and School Year Trainings in 2012–2013

Most administrators (78%) were trained during the summer. Those who were trained during the school year had largely similar characteristics and experiences to those who completed training in the summer (see Tables B13b, B15b, and B26b). Both groups held similar views of the TGDC, the initial training, their job efficacy, and skills that make an effective teacher. In the summer, however, larger proportions of principals and assistant principals attended training, while non–building administrative staff were more likely to attend training during the school year (see Table 1).

A more notable difference was that administrators trained during the school year reported feeling more comfortable with the TLF. This difference may have occurred for many reasons. Trainers may have felt more comfortable after the first round of trainings in the summer. The quality of the training may actually have improved, due to changes made over time. Changes may have been made over time. Administrators with different, nonbuilding roles may have viewed training differently.

Administrators who attended training during the school year were also more likely to say that the online system, MyPGS, made the TGDC easy to use. Again, we can only speculate about the source of this difference. Many changes were made to MyPGS over time, and trainers undoubtedly grew increasingly familiar with it over time. Together, these changes and increasing familiarity may have contributed to more positive views of MyPGS.

Training Changes for 2013–2014

During summer 2013, LAUSD offered two trainings: (a) a mandatory 5-day training for administrators who had not previously attended training and (b) an optional 1-day refresher course. Both trainings featured an increased emphasis on the role of coaching in the TGDC.

Trainers conveyed that it was not sufficient for administrators to become certified; they must also be able to have conversations with teachers that result in improved practice. The refresher course, taught by two teaching and learning coordinators (TLCs), focused on having coaching conversations with teachers and on new improvements to MyPGS.

TLS provided the 5-day initial training for new administrators. This training reflected the increased emphasis on coaching, with a full day devoted to strategies for coaching conversations. Members of the research team attended the new training and noticed several minor revisions. In contrast to the summer 2012 training, in 2013, trainers moved more quickly into the content of the TLF. Whereas the previous year's training began with a conversation about administrators' hopes and fears for the new evaluation system, this year's training jumped right to the rationale for selecting the TLF as the district's observation instrument. The TLS trainer defined each of the TLF's focus elements and grounded the definitions in examples of video-recorded lessons. Each video was followed by a targeted conversation about why instruction was rated at a certain level. Whereas rationales for rating specific elements (e.g., "What makes this Developing on Quality/Purpose of Questions, and not Effective?") were not emphasized during the summer 2012 training, these discussions and justifications were more extensive in 2013. The new iteration of the initial administrator training also emphasized adhering to the language of the rubric as a precaution against thinking of teachers in categorical terms (e.g., "this person is a 2"). Finally, the trainer pointed out that most teaching practice will not be distinguished. One aspect of the training that was the same in 2013 was the strong emphasis on taking objective notes.

How Did Administrators Perform on the Certification Assessment?

Summary of Findings

At the conclusion of their initial trainings, most administrators attained the level of preliminary certification, and a small number scored at full certification. The greatest number of administrators became certified on objectivity, while more struggled with representation and accuracy. The components of certification (objectivity, alignment, representation, and accuracy) were both related to and distinct from one another. The certification component with which administrators struggled most—accuracy—was unrelated to the other three components. When rating the same video-recorded lesson, administrators did not reliably agree with master raters or with each other, a finding we describe in more detail in what follows.

What Aspects of the Teaching and Learning Framework Are Most Challenging?

For teachers and administrators to improve instruction, they must have accurate and reliable information on which to base their actions. LAUSD has worked to promote accuracy and reliability by requiring all observers to attain certification. Certification serves as a check on what was learned in training and identifies disagreement among observers about evidence of high-quality teaching. Certification assessment results from each training session were analyzed to determine how well administrators agreed with one another and with the master raters. We considered agreement on certification ratings from a number of vantage points. First, we looked at correlations among the four components of certification (objectivity, alignment, representation, and accuracy). We also calculated agreement between certified administrators and master raters as well as administrators' agreement with one another on the four standards of the TLF that appeared on the certification test. We generated results for the whole group of administrators by job title and by the grade level of students served.

TLS created its four certification components from both qualitative and quantitative measures. The qualitative measures focused on observers' assessment of objectivity, alignment, and representation from the scoring rubric. *Objectivity* was defined by bias-free evidence; *alignment* referred to the proper placement of evidence in the standards and elements of the TLF; and *representation* indicated that sufficient evidence from raters' scripting that support the scores assigned was noted. Quantitative measures were derived by comparing administrators' ratings to those of a master rater or an expert observer. Three different subscores made up this measure: *accuracy* on discrete items (the number of exact matches with the master rater), the *score differential* (the difference between average ratings for each standard), and a *volatility index* (the difference between discrete items plus the difference between element-level ratings of the master rater, adjusted for the cohort group average).

Overall Certification Performance

Table 4 presents overall certification rates by component (objectivity, alignment, representation, and accuracy) for each of two certification videos: *Miller* ($N = 785$) and *Infelice* ($N = 324$). The difference in sample sizes reflects the difference in participants trained in the summer (*Miller*) and school year (*Infelice*). Regardless of video, most administrators (83% of those who watched and rated *Miller* and 81% who rated *Infelice*) attained certification at the level of preliminarily certified. Few administrators (11%) failed to attain certification, and even

fewer (6% and 8%, respectively) attained full certification. In terms of the four certification components, the greatest number of administrators certified on objectivity or collecting bias-free evidence. This is not surprising given that taking objective notes was a primary focus of training throughout 2012–2013. The components that administrators struggled most with were representation and accuracy; about one third of administrators (31% and 35%, respectively) failed to become certified on accuracy. Finally, although there were small differences in certification rates by video (i.e., in accuracy and alignment), we do not have the data to parse out whether these differences were due to the relative difficulty of the two videos or to differences between the population of administrators completing training during the summer and school year, respectively.

Table 4. Miller and Infelice Certification Rates (Overall and by Component)

Certification	Overall		Objectivity		Alignment		Representation		Accuracy	
	M	I	M	I	M	I	M	I	M	I
Certified with distinction	0	0	<1	0	<1	0	0	0	6	1
Certified	6	8	76	77	41	50	30	26	26	34
Preliminarily certified	83	81	18	19	45	42	56	56	37	30
Not yet certified	11	11	5	4	14	7	13	17	31	35

Note. M = Miller ($N = 785$); I = Infelice ($N = 324$). Responses listed as percentages, with frequencies in parentheses.

Given that administrators were least likely to become certified on the accuracy component, we also tested whether certification rates would have changed if LAUSD had relied on only one of the three accuracy categories (i.e., discrete item accuracy, score differential, or volatility index) rather than a composite measure. We replaced administrators' overall accuracy scores with their scores for each accuracy category using Summer 2012 data (see Table 5). The results of this analysis indicated that the composite measure used by LAUSD was slightly more generous; that is, a greater percentage of administrators scored at or above preliminarily certified than in any single category. Of the three accuracy categories, only the volatility index seemed to change certification outcomes, with only 1% of administrators attaining full certification and almost a quarter (21%) not meeting the certification standard.

We also investigated the degree to which certification components were related to one another and found small to moderate correlations (see Table B4). None of the certification components, however, were correlated more strongly than .40. The small to moderate correlations support the idea that the four components of the certification assessment were

related but also distinct, measuring somewhat different underlying constructs. The component that most administrators struggled with, accuracy, was not related to the other components. When rating the same video, administrators did not agree with master raters or with one another often enough to establish an acceptable level of reliability.

Table 5. Certification Rates by Quantitative Measures (Existing Composite and Accuracy Categories)

Certification Rate	Accuracy composite	Discrete item accuracy	Score differential	Volatility index
Certified with distinction	0 (0)	0 (0)	0 (0)	0 (0)
Certified	6 (49)	5 (41)	6 (46)	1 (7)
Preliminarily certified	83 (650)	82 (642)	80 (628)	78 (616)
Not yet certified	11 (86)	13 (102)	14 (111)	21 (162)

Note. Summer 2012 ($N = 785$). Responses listed as percentages, with frequencies in parentheses.

In addition, we examined whether certain standards were more challenging to rate than others (see Tables C5–C7). The highest levels of interrater agreement (66%–72% exact match with master raters) were associated with Standard 2 (Designing Coherent Instruction), which focuses on behavior management, organization, and the classroom environment. The lowest levels of interrater agreement were found for Standard 1 (Planning and Preparation, 17%–65% match) and Standard 5 (Professional Growth, 43%). Specifically, administrators tended to rate teachers at higher levels than master raters did. On the remaining standard, Standard 3 (Standards-Based Learning Activities, 10%–67% match), some of administrators' ratings were higher than master raters, whereas others were lower.

Finally, we examined variation in administrators' certification outcomes by job role and grade level. In terms of job roles, principals agreed more often with master raters than did assistant principals across all TLF standards (see Tables C5–C7); these differences were large on Standard 5. When we looked at the elements that principals and assistant principals rated most and least similar to master raters, we identified three elements with great disparities in ratings for both groups: design of formative assessments (17% match), communicating the purpose of the lesson (20% match), and assessment criteria (10% match). Two of these elements are organized under Standard 3. Principals agreed with master raters most often on monitoring and responding to student behavior (75%); management of routines, procedures, and transitions (75%); and teacher interactions with students (73%). All three of these elements are organized under Standard 2. Assistant principals agreed with master raters most often on ratings of expectations

for learning and achievement (73%), discussion techniques (69%), monitoring and responding to student behavior (68%), and creating purposeful instructional groups (68%). Two of these four elements are organized under Standard 2, which also had the highest levels of reliability overall.

We found a similar pattern of administrator agreement with master raters when we parsed them by the grade level of their schools (see Table B7). Across grade levels, more administrators agreed with master raters for elements of Standard 2. Administrators who worked in schools at different grade levels varied slightly on which elements they rated most accurately; the elements for which agreement with master raters was lowest, however, were the same across grade levels: design of formative assessments (16%–18%), communicating the purpose of the lesson (15%–22%), and assessment criteria (8%–11%).

Certification is only one aspect of quality control in a teacher observation system. It is important to the validity of TGDC that it does not matter which administrator assigned observation scores to a teacher. Therefore, we discuss in more detail administrators' accuracy and reliability using TLF to rate teaching. To assess the degree to which a teacher's score would not be influenced by the rater, we considered the agreement rates of administrators with one another. To calculate agreement between raters, we randomly assigned raters to pairs and then looked at, for every element, the percentage of pairs that exactly matched each other; scores were then averaged by standard. These statistics help us understand the degree to which one administrator would have given the same lesson a similar score as another administrator. If agreement rates are high, this lessens the chance that a teacher's score will depend on the observer.

One can also look at agreement rates between administrators and master raters; this is similar to what is reported for the accuracy component of certification in the preceding section. For every element, we simply calculated the percentage of administrators whose scores matched the master raters exactly. We also calculated the average deviation from the true score (i.e., the score created by the master rater) to determine whether administrators were generally over- or underestimating the teacher's performance on a given element. The more administrators and master raters agree, the more likely the administrators are implementing the protocol accurately.

Though there is little technical scholarship that documents what an acceptable level of agreement might be for observations, we use 80% agreement between master raters and administrators as a rough standard. This means that 20% of the time, it would be acceptable for

master raters and administrators not to agree on a score. Certification data suggest that administrators generally did not agree with master raters at acceptable levels (Table B5). In all four standards that were scored for certification—Standard 1: Planning and Preparation, Standard 2: Designing Coherent Instruction, Standard 3: Standards-Based Learning Activities, and Standard 5: Professional Growth—administrators did not match master raters at levels at or above 80%. They also did not match one another at levels at or above 80%. Administrators demonstrated the most accurate use of the scoring protocol in Standard 2, agreeing with master raters 69% of the time. The other three areas of the TLF had similar levels of agreement with master raters (42%–48%). The results were similar for the agreement between administrators, although on Standards 2 and 5, intra-administrator exact agreement was considerably worse than agreement with master observers (52% vs. 69% and 32% vs. 43%).

The mean deviation of scores tells which direction and how far off administrators were from master scores. At the standard level, administrators assigned scores that were roughly a quarter of a point higher than master scores (.23–.24) for Standards 1 and 5. Administrators on average assigned scores that were just over a tenth (–.11) of a point lower than master scores for Standard 2. For Standard 3, administrators assigned scores that were about one tenth (.12) of a point higher than master scores.

Administrators most closely matched master raters on elements that judge the quality of behavior, climate, and management of the classroom. The four elements with the highest agreement between administrators and master scores concerned monitoring student behavior (72% agreement), managing the classroom learning environment (70% agreement), interacting with students (69% agreement), and setting expectations for learning (69% agreement).

What Predicted Certification and Training Satisfaction?

Summary of Findings

The strongest predictor of training satisfaction was the administrator's belief in the value of teacher evaluation. In contrast, positive perceptions of job manageability appeared to undermine satisfaction. There was also a positive association between administrators' initial expectations for training and the degree to which those expectations were met by the training. There were no significant differences in training satisfaction by the specific job an individual held (e.g., central office administrator, assistant principal) nor by the time of training (i.e., summer or school year).

In the model predicting overall certification outcomes, only one variable was associated with successfully attaining certification: importance of teacher evaluation data. Interestingly, it was a small and negative association; administrators who valued teacher evaluation data were slightly less likely to attain certification. Although there were stronger associations when examining the association between the pre- and posttraining measures and the certification subscores, the relationships were counterintuitive. Results suggested that principals were more likely to attain certification on accuracy than other administrators.

Background

In this section, we describe our efforts to explore associations between training survey data and two outcomes: (a) satisfaction with training and (b) certification success. On their own, each of the outcomes is likely valuable to LAUSD in planning for future training sessions. Certification is a prerequisite for administrators to conduct observations of their teachers, whereas satisfaction data can guide efforts to improve certain aspects of the training over time. If, for example, administrators lack confidence in aligning evidence, training satisfaction data can identify this area for growth. It was also important to us to examine whether we could predict these outcomes using data on administrators' characteristics, beliefs, or experiences during training. Using survey data collected before, during, and after the training sessions, we conducted a series of regression analyses predicting each outcome. Doing so allowed us to make comparisons between administrators' perceptions of training (i.e., the posttraining survey outcomes) and their actual training performance (i.e., the certification results).

Predicting Administrator Satisfaction With Training

Outcome measures. Drawing on the posttraining survey, we used three indicators of training satisfaction: (a) participants' self-assessment of learning in training, (b) their confidence in their ability to reliably and accurately conduct observations, and (c) their assessment of whether the training met their expectations. Each of these three indicators was a composite score made up of several survey items. Composite scores were calculated using the means of individuals' survey items related to the satisfaction indicators, which was on a 4-point scale ranging from 1 (*not a lot or disagree*) to 4 (*a lot or agree*). Specifically, learning in training was calculated using posttraining survey items 1a–1i (mean = 3.42, $SD = .48$, $\alpha = .85$); confidence in conducting observations was calculated from survey items 5a–5i (mean = 3.34, $SD = .45$, $\alpha =$

.87); and training met expectations was calculated from survey items 2a–2d (mean = 3.42, $SD = 0.60$, $\alpha = .87$). All individual items ranged from 1 (*not a lot or disagree*) to 4 (*a lot or agree*).

Predictors. We used two categories of independent variables when predicting training satisfaction: (a) administrator job characteristics and (b) administrator beliefs and expectations for training. We focused on two administrator job characteristics—the instructional level of their positions (i.e., elementary, middle, high, other) and their job role (i.e., principal, assistant principal, other). We acknowledge that there were several other categories of job role not covered in these categories, such as instructional director, but group sizes for these variables were not large enough to be included in our regression analyses. Each job characteristic variable was treated dichotomously, for example, *principal* = 1, *not principal* = 0.

There were a series of pretraining survey composite measures representing administrator beliefs and expectations for training. These variables are summarized in Table 6.

Table 6. Pretraining Survey Measures Used to Predict Training Satisfaction

Survey measure	Presurvey items	Mean response (1–4)		
		Overall	Summer training	School year training
Importance of student growth data in teacher evaluation	6f	3.34 (0.69)	3.31 (0.70)	3.46 (0.64)
Importance of teacher evaluation data	7a–7g	3.12 (0.66)	3.08 (0.67)	3.27 (0.58)
Perception of job manageability	11a–11d	3.15 (0.46)	3.15 (0.46)	3.16 (0.45)
High expectations for observation training	5a–5i	3.67 (0.44)	3.65 (0.45)	3.74 (0.36)
Belief in student-centered instruction	9f–9j	3.82 (0.29)	3.81 (0.30)	3.86 (0.23)

Note. Standard deviations are in parentheses. To reduce text in the table, we have changed the wording of these variables from the original items used in the surveys.

As with the outcome measures, we used mean composites for all belief measures, with the exception of the importance of student growth data in teacher evaluation measure, which was a single survey item.

Other control variables. In addition to administrator characteristics and beliefs items, we also controlled for week of training, given the potential variation in training experience (and thus satisfaction) across training weeks. Similarly, we anticipated that training might vary by training site (as LAUSD trained administrators at four different sites). Thus we used robust standard errors in our regression analyses to account for the clustering of administrators within training sites. Note that we considered the use of hierarchical linear modeling to address this clustering, but the variation attributable to training site was negligible. We did not see significant differences in satisfaction across the four training sites.

Results. Across all three measures of training satisfaction, the greatest predictor of satisfaction was the administrator's belief in the importance of teacher evaluation data. As our analyses suggest, when administrators believed in the importance of teacher evaluation data, they were more likely to be satisfied with the TGDC training. Our results also indicated that positive perceptions of job manageability were predictive of administrators' confidence in conducting observations. There was also a positive association between administrators' initial expectations for training and the degree to which those expectations were met by the training. There were no significant differences in training satisfaction by the specific job an individual held (e.g., central office administrator, assistant principal). Finally, administrators trained during the 2012–2013 school year were no more or less satisfied with their training than those trained during summer 2012. Results are summarized in Table 7.

Predicting Certification Outcomes

Outcome measures. Overall certification was the most relevant outcome for this analysis. However, the variation in the administrators' overall certification scores was highly restricted. Most administrators (83%) were preliminarily certified, and only 10% were fully certified. Consequently, we decided to use the four certification measures developed by TLS—alignment, objectivity, representation, and accuracy—to identify more variation in certification success at the level of these constituent measures. Many administrators achieved certification in a single area, making it possible for us to treat the individual measures as outcomes in our analyses. In all cases, we treated the outcome measures dichotomously, where 1 = *certified* (C) and 0 = *not yet certified* (NYC) or *preliminarily certified* (PC). We double-checked this approach by treating both preliminary and full certification as certified, and this did not change our findings.

Table 7. Regression Coefficients for Predicting Training Satisfaction Using Pretraining Measures of Beliefs and Expectations (Standard Errors)

Belief/Expectation	Model 1: Assessment of learning	Model 2: Confidence in conducting observations	Model 3: Training met expectations
Pretraining beliefs/expectations			
Importance of student growth data in teacher evaluation	.05 (.05)	-.01 (.04)	.02 (.06)
Importance of teacher evaluation data	.23*** (.06)	.14** (.06)	.17*** (.05)
High expectations for observation training	.10* (.05)	-.02 (.06)	.22** (.08)
Perception of job manageability	.05 (.05)	.11*** (.02)	-.03 (.04)
Belief in student-centered instruction	-.14* (.08)	.04 (.06)	.06 (.05)
Administrator characteristics			
Elementary	-.11 (.09)	-.03 (.05)	.03 (.05)
Middle	-.09 (.07)	-.09* (.05)	-.04 (.04)
High	.03 (.10)	.09 (.05)	.08 (.05)
Principal	-.04 (.09)	-.07 (.10)	-.11 (.10)
Assistant principal	.06 (.11)	-.07 (.09)	-.17 (.13)
Training type: summer (vs. school year)	-.06 (.06)	-.04 (.07)	-.03 (.08)

Note. $N = 408$. Standard errors are in parentheses. To account for shared experiences within training sites, we included robust standard errors.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Predictors. As in the previous analysis, we used information on administrators' job characteristics and their beliefs and expectations related to training. From the pretraining survey, we used three measures that we hypothesized predicted certification success: (a) importance of teacher evaluation data, (b) high expectations for observation training, and (c) perception of job manageability. From the posttraining survey, we used the three outcomes related to training satisfaction: (a) assessment of learning, (b) training met expectations, and (c) confidence in ability to conduct observations. Information on instructional level and job title is also included.

Results. We ran a series of ordered logit models in which we estimated the likelihood of administrators scoring at the NYC, PC, or C level. Each model controlled for a number of administrator role characteristics and whether administrators were trained in the summer or during the school year. We used robust standard errors to account for similarities in experiences within training sites. Results from these analyses are presented in Table 8. In the model predicting overall certification, only one variable was associated with certification: importance of teacher evaluation data. Interestingly, it was a small and negative association; administrators who more valued teacher evaluation data were slightly less likely to become certified.

Table 8. Regression Coefficients Predicting Certification Using Pre- and Posttraining Measures of Beliefs and Expectations (Standard Errors)

Belief/Expectation	Model 1: Overall certification	Model 2: Objectivity	Model 3: Alignment	Model 4: Representation	Model 5: Accuracy
Presurvey measures					
Importance of teacher evaluation data	-.18* (.11)	.20 (.24)	-.10 (.08)	-.11 (.26)	-.04 (.12)
High expectations for observation training	.37 (.35)	-.23 (.35)	.16 (.31)	.13 (.21)	-.08 (.26)
Perception of job manageability	-.15 (.27)	.15 (.22)	-.17 (.23)	-.23** (.09)	-.34 (.40)
Postsurvey measures					
Assessment of learning	-.25(.21)	-.72*** (.27)	-.14 (.11)	-.06 (.20)	-.30* (.18)
Training met expectations	.20(.24)	.27 (.20)	.39 (.24)	.39** (.18)	-.10 (.25)
Confidence to conduct observations	.27 (.48)	.27 (.24)	-.16 (.34)	.26 (.28)	-.02 (.15)
Administrator characteristics					
Elementary	.49 (.33)	.30 (.54)	.040 (.21)	.07 (.37)	-.03 (.35)
Middle	-.15 (.38)	-.63 (.49)	.61** (.29)	.11 (.31)	.06 (.31)
High	.43 (.43)	.61 (.59)	-.53 (.43)	-.14 (.32)	-.35 (.54)
Principal	.02 (.41)	.35 (.57)	.21 (.31)	-.28 (.45)	.82* (.47)
Assistant principal	-.41 (.80)	.53 (.73)	-.23 (.49)	.03 (.18)	.47 (.66)
Training type: summer (vs. school year)	.89 (.23)	.42 (.93)	.13 (.27)	1.48 (.42)	-.04 (.29)

Note. $N = 349$. All certification variables coded as follows: 3 = certified with distinction or certified; 2 = preliminarily certified; 1 = not yet certified.

* $p < .10$. ** $p < .05$. *** $p < .01$.

In comparison to the model predicting overall certification, there were stronger associations between the pre- and posttraining survey measures and certification component scores. Of the pretraining survey measures, the only one predictive of any certification component was “perception of job manageability.” However, the results related to job manageability were counterintuitive: The negative value indicates that administrators who perceived their jobs as more manageable were less likely to become certified on representation. Similar negative associations were found on the posttraining survey. Administrators who felt that they learned a lot during training were less likely to become certified on both objectivity and on accuracy. This finding may be the result of administrator familiarity with the TLF; specifically, administrators who had views of teaching that aligned with the TLF may have had less to learn during training and thus felt that they did not learn as much. These administrators may have been more likely to become certified due to their background knowledge. However, this rationale is speculative. Of the remaining postsurvey measures, it appeared that there was a positive and

significant association between whether the training met participants' expectations and their likelihood of becoming certified on representation. With regard to administrator characteristics, middle school administrators were more likely to become certified on alignment. Similar to our descriptive analyses, principals were more likely to become certified on accuracy than other administrators were.

How Did the 2012–2013 Observations With Teachers Go?

Summary of Findings

Administrators had a range of experiences with the TGDC. There was wide variation in the number of formal and informal observations conducted as well as the time it took to complete them. Regarding formal and informal observations, focus administrators believe that informal observations are helpful for supporting development and for giving information that is useful in managing their staff. They reported needing both informal and formal observations to support teacher development. When reflecting on their experiences in 2012–2013 using the TGDC, administrators raised two major concerns: the time to conduct formal observations and the technical aspects of the MyPGS platform. Additionally, some administrators reported that it was difficult to find a focus teacher to work with in the first place, citing teachers' substantive concerns with participating. Despite these concerns, administrators viewed the TLF itself and the yearlong TGDC experiences favorably. The most requested training needs for 2013–2014 were additional support for learning how to align evidence and managing the transition to full implementation.

Background

We collected survey and interview data to evaluate the 2012–2013 observations with teachers. Specifically, we collected data on all administrators through an end-of-year (EOY) survey and collected more in-depth data from focus administrators via Spring 2013 interviews and the bimonthly surveys.

Finding a Focus Teacher

Finding a focus teacher is critical to administrators' abilities to learn how to use the TGDC—without a teacher to practice with, the administrator has no hands-on opportunity to learn. On the basis of focus administrators' responses, we learned that teachers came to

participate in the 2012–2013 TGDC rollout in one of three major ways. They were asked by their administrator (56%), they volunteered after volunteers were solicited (34%), or they approached their administrator and said they were interested (9%). Just over 1 in 10 of the administrators who responded to the EOY survey reported that it was difficult to find a focus teacher to work with. Most concerning, among all administrators surveyed at the end of the 2012–2013 school year, 26% reported that they were unable to work with a focus teacher for this year (see Table 9).

Table 9. Administrators’ Experience With Focus Teachers in 2012–2013, End-of-Year Survey Sample

Statement	Percentage	Count
I was not able to work with a focus teacher this year	25.6	100
I conducted formal observations with my focus teacher	67.5	264
I conducted informal observations with my focus teacher	54.7	214
I did not conduct any observations with my focus teacher	5.9	23
It was difficult to find a focus teacher to work with this year	12.0	47

Note. $N = 398$. Percentages sum to greater than 100% because participants were allowed to select multiple items.

Based on interview data with the focus administrators, the challenges associated with finding a focus teacher to work with came from many places—teachers’ substantive concerns, logistical challenges, and administrator workload challenges. Some challenges came from teacher concerns—teachers did not volunteer because they objected to the TGDC policy; some administrators said that the union discouraged teachers from participating. Other challenges were logistical. Administrators changed jobs during the year, which made it difficult to conduct the observations and conversations. Focus teachers went out on maternity leave and could not participate, and it was too late for the administrator to find a new teacher with whom to work. Often scheduling conflicts made it difficult to conduct a full observation cycle. For example, principals reported that they would hold a preobservation conference with a teacher, and then on the way to observe the lesson an urgent matter would arise, making it impossible for the principal to observe the lesson discussed in the preobservation conference. Particularly for secondary teachers, where the lesson content changes rapidly, this would mean that a new preobservation conference would need to be scheduled. Thus some administrators never commenced an observation cycle at all; others started but were not able to complete one full cycle with any teacher. And some were related to workload; some administrators just did not have time, given their other responsibilities, to find a teacher with whom to work. In any case, a quarter of all administrators who will be responsible for conducting observations and ratings for

multiple teachers in the coming year have not yet conducted one full cycle in a relatively low-stakes environment.

Experiences With Formal and Informal Observations

As described previously, LAUSD administrators were tasked with conducting two formal observations with their focus teachers. In addition, they were encouraged to schedule additional informal observations, which were less structured, and scores from them were not uploaded to the online scoring system. Administrators who responded to the EOY survey had a range of experiences with the TGDC. As mentioned previously, just over one quarter of administrators did not work with a focus teacher in 2012–2013. Of those who did work with a focus teacher, two thirds conducted at least one formal observation, about 55% conducted at least one informal observation, and 6% did not do any observations with a teacher. Administrators reported completing an average of 2.0 formal observations and 4.9 informal observations (see Table 10). The average formal observation took just longer than 10 hours. Focus administrators reported that informal observations took an average of 2 hours.

Table 10. Statistics on Formal and Informal Observations Administrators Did With Focus Teachers in 2012–2013, End-of-Year Survey Sample

Observation	Mean	SD
Formal observation(s)	2.0	0.9
Informal observation(s)	4.9	7.9

Note. $N = 398$.

There was wide variation in the number of formal and informal observations conducted as well as in how long those observations took to complete. As an example, administrators reported a time range for formal observations between 1 and 120 hours—or almost 15 school days. The validity of these estimates is unclear; it seems unlikely that administrators would be able to spend 120 hours on one formal cycle of the TGDC.

Because informal observations were conducted frequently by administrators, focus administrators were asked about how they used informal observations and the ways in which formal and informal observations differed (see Table 11). Roughly 20% of administrators used formal and informal observations in similar ways. Administrators believed that informal observations were helpful in supporting development and providing information that is useful for staff management. In the mid-year survey, we asked focal administrators about formal versus

informal observations. Focus administrators agreed that they needed both informal and formal observations to support teacher development. Three quarters (75%) of focus administrators agreed or agreed somewhat that informal observations were helpful but cannot replace formal observations. Almost all (97%) reported that informal observations were helpful in managing their staff, and a large majority of administrators (92%) agreed or agreed somewhat that informal observations gave them adequate information to support teacher development.

Table 11. Focus Administrator Agreement Regarding Informal Observation, Mid-Year Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
Informal observations provide information that is helpful to me in managing my staff.	0.0% (0)	2.8% (1)	27.8% (10)	69.4% (25)
Informal observations give me enough information to support the teacher's development.	2.8% (1)	5.6% (2)	47.2% (17)	44.4% (16)
Informal observations are helpful but cannot replace formal observations.	8.3% (3)	16.7% (6)	30.6% (11)	44.4% (16)
I use informal observations differently than I use formal observations.	5.6% (2)	13.9% (5)	44.4% (16)	36.1% (13)
I need both informal and formal observations to support teacher development.	0.0% (0)	5.6% (2)	25.0% (9)	69.4% (25)

Note. $N = 36$. Frequencies are in parentheses.

One example of how informal and formal observations might be used differently came up in interviews with administrators. A number of them mentioned that some TGDC requirements did not apply easily to a small number of cases of teachers with a history of unsatisfactory teaching. One assistant principal explained that if she and her principal were deeply concerned about a teacher's performance, they did not go into the teacher's classroom for only two formal observations, as the TGDC requires. Instead, they had to go into the classroom almost monthly. Two other principals shared the same concerns about the application of the TGDC's requirements to this group of teachers. Each of the more regular visits that these types of situations demanded also often required formal (rather than more focused informal) observations. Administrators agreed that it would be difficult to do that many formal observations for more than one or two teachers per year.

At the end of the year, when reflecting on their experiences in 2012–2013 using the TGDC, administrators expressed a range of opinions. The time it takes to conduct formal observations is a serious concern for most administrators. Based on the responses of 398 administrators, 75% reported that they felt formal observations took too much time, 18%

reported they took an appropriate amount of time, and 7% reported they took too little time. As detailed elsewhere, the issue of time to conduct formal observations is the most serious concern in the implementation of TGDC. One focus rater described the following:

I do believe that the greatest drawback to the TGDC is the amount of time needed to complete the process. Single site administrators would be hard-pressed to perform all the operational tasks required and then complete full TGDC cycles.

Administrators felt that there are many important aspects of their jobs and that they must balance TGDC with other obligations. Administrators' comments about other important and time-consuming aspects of their jobs are listed in Tables B6b and B6c. A smaller proportion of administrators raised a second concern about the technical aspects of the MyPGS platform.

Despite concerns about time and the technology platform, there was a high level of administrator buy-in to the rubric itself. As we saw in other surveys, administrators viewed the TLF favorably. Table A3 shows that almost all (99%) agreed or agreed somewhat that the TLF covers important domains of teaching and is consistent with their own views of teaching. Ninety-four percent (94%) agreed or agreed somewhat that the TLF helped them think more critically about teaching, and a similar portion (96%) agreed or agreed somewhat that the TLF is a valuable tool for talking to their teachers about instruction. Most administrators (89%) were hopeful that the TLF will improve the quality of teaching in their schools and meet the learning needs of their students and teachers.

The administrators also expressed favorable opinions of the yearlong TGDC experiences. Most administrators (75%) who returned the EOY survey said they agreed or somewhat agreed that the training prepared them for their work in 2013–2014 (see Table 12). Administrators agreed or somewhat agreed that they were able to do critical components of the TGDC: collect objective evidence (92%), conduct preobservation conferences (94%), conduct postobservation conferences (94%), understand the differences between performance levels on the TLF (95%), and understand the TLF's components and elements (96%). One administrator summarized the value of the TGDC as follows: "TGDC has been an excellent professional growth experience, and I am grateful for the new skills I have gathered throughout the process." An area where some administrators (24% disagreed or somewhat disagreed) noted concerns was their confidence in their ability to use the online observation tool, MyPGS. A small minority (14% disagreed or

somewhat disagreed) also expressed concern about their ability to align evidence properly. A wide range of administrators' reflections on the TGDC practice year is listed in Table A40.

Table 12. Administrators' Views of Training Outcomes, End of-Year Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I feel confident in my ability to use MyPGS, the online observation tool.	6.3% (25)	17.8% (71)	49.5% (197)	26.4% (105)
I feel confident in being able to collect objective evidence of instruction.	1.5% (6)	6.4% (25)	46.6% (183)	45.5% (179)
I feel confident in my ability to align evidence.	2.5% (10)	11.1% (44)	52.9% (210)	33.5% (133)
I understand the standards, components, and elements in the LAUSD Teaching and Learning rubrics.	1.0% (4)	3.3% (13)	38.2% (151)	57.5% (227)
I understand the differences between the levels of performance (highly effective, effective, developing, and ineffective) in the Teaching and Learning rubrics.	1.0% (4)	4.3% (17)	37.8% (150)	56.9% (226)
I feel confident conducting preobservation meetings with teachers.	0.3% (1)	5.8% (23)	36.2% (144)	57.8% (230)
I feel confident conducting postobservation meetings with teachers.	0.5% (2)	5.3% (21)	36.1% (143)	58.1% (230)
LAUSD's training prepared me adequately for my observation work this year.	6.1% (24)	18.8% (74)	53.0% (209)	22.1% (87)

Note. $N = 398$. Frequencies are in parentheses. LAUSD = Los Angeles Unified School District.

Reflections on Future Training Needs

Focus administrators suggested that working with teachers in 2012–2013 was a positive experience. One administrator commented:

The discussions with the teacher I was observing were an integral part of shortening my learning curve. The frequent instructional rounds using the TLF with my instructional director and support team were beneficial in using and becoming more familiar with the TLF.

But administrators also recognized that the teachers they were working with were different than many of the teachers in their buildings. In general, they were strong teachers, and they were open to the TGDC. Administrators understood this would not necessarily be the case next year. In the words of one principal, “I believe next year will be a real test of the full implementation as it will be replacing actual teacher evaluation.”

In part because of the limited rollout and moderate number of actual experiences working with the TLF, we asked focus administrators to identify areas with which they most needed assistance in the 2013–2014 school year. Of all the statements in Table 13, administrators were

asked to rank the top three. By far the most popular needs were additional support learning how to align evidence and managing the transition to full implementation, with about 60% and 65%, respectively, of administrators ranking those needs in the top three. There were four other needs that 30%–32% of administrators ranked in their top three: strategies for working with teachers who can improve, scoring preobservation plans and artifacts, understanding the differences between performance levels, and matching master raters' scores. Just three administrators said that support learning about the core concepts of the TLF was one of their top three professional learning needs.

Table 13. Focus Administrators' Rankings of Professional Learning Needs

Learning Need	Rank 1	Rank 2	Rank 3	Total
Managing the transition from the pilot to full implementation of the TGDC observations	27.0% (10)	5.4% (2)	32.4% (12)	64.9% (24)
Aligning evidence	27.0% (10)	21.6% (8)	10.8% (4)	59.5% (22)
Strategies for working with teachers who I believe can improve but may be somewhat resistant	13.5% (5)	10.8% (4)	8.1% (3)	32.4% (12)
Scoring teachers' lesson plans and other preobservation materials	5.4% (2)	10.8% (4)	16.2% (6)	32.4% (12)
Understanding the differences among levels on the rubric (highly effective, effective, developing, and ineffective)	10.8% (4)	5.4% (2)	13.5% (5)	29.7% (11)
Matching the master raters' scores on training videos	2.7% (1)	24.3% (9)	2.7% (1)	29.7% (11)
Giving teachers feedback on their formal observations	5.4% (2)	13.5% (5)	5.4% (2)	24.3% (9)
Understanding the whole TGDC observation process	5.4% (2)	5.4% (2)	8.1% (3)	18.9% (7)
Understanding the core concepts in the Teaching and Learning Framework	2.7% (1)	2.7% (1)	2.7% (1)	8.1% (3)

Note. $N = 37$. Frequencies are in parentheses. TGDC = Teacher Growth and Development Cycle.

How Did Administrators' Views Change Over the School Year?

Summary of Findings

In general, administrators' views did not change dramatically over the year of training and practice on the TGDC. We considered administrators' views in three areas: their support for the multiple goals of the TGDC system, their views of the TLF, and their efficacy in their job roles. Administrators were most optimistic that the TGDC can support teacher improvement and development. They were least optimistic that it will support teacher dismissal decisions or increase their ability to spend more time on teaching and learning. Administrators' beliefs that the TGDC system could shape many human capital management decisions remained steady on

four goals (teacher improvement/development, school improvement, teacher dismissal, and public accountability) and shifted down on three (the ability of the system to identify or reward strong teachers, create a common vision of teaching, and improve the district). Notably, administrators' views of the TLF did not change. They left their week of training and the practice year supportive of the quality and relevance of the TLF. The one exception to this pattern was that over the school year, administrators became less confident in their ability to navigate MyPGS. Finally, administrators' views of their efficacy and focus were mixed. Their focus on decision-making using their school improvement plan increased slightly. Their beliefs in their abilities to manage their jobs and control operational decisions decreased.

Administrators' Support of Teacher Growth and Development Cycle Human Capital Goals

The district is trying to provide data support to administrators as they make human capital decisions such as the content of professional development sessions and promotions to coaching or leadership positions. As described in the district Teacher Incentive Fund (TIF) grant proposal, there are many human capital management goals of the TGDC (U.S. Department of Education, 2012). It is important to understand the degree to which the individuals charged with carrying out these reforms believe the data are useful for the range of purposes the district is intending.

We think of the changes in administrators' views about the purposes of the TGDC in two time blocks. The first block is the daily intense time devoted to the 5-day initial training. The second block is the school year in which administrators were trying out the TGDC with real teachers and continuing their learning. We tested for significant differences between administrators' views of the purpose for TGDC data before and after the initial training using dependent *t*-tests. We found that, with one exception, there was no change in administrators' ratings of the usefulness of the various purposes listed (see Table 14). Only the purpose, identifying or rewarding strong teachers, increased from pre- to posttraining, with the pre-post difference significant at $p < .01$. Both before and after the initial training, administrators were most optimistic that the TGDC could be used to support teacher improvement and development and create a common vision of quality teaching (Tables B9 and B28). They were least optimistic that the TGDC could support teacher dismissal and district improvement.

Table 14. Administrators' Mean Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes Before and After Training

View	N	Before	After	<i>p</i> value
Identifying or rewarding strong teachers	412	2.95 (0.83)	3.15 (0.80)	0.00***
Teacher improvement/development	410	3.36 (0.72)	3.33 (0.70)	0.53
School improvement	410	3.29 (0.75)	3.24 (0.73)	0.37
Teacher dismissal	408	2.87 (0.94)	2.87 (0.95)	0.89
Creating a common vision of excellent teaching	413	3.36 (0.75)	3.39 (0.71)	0.43
District improvement	412	3.04 (0.84)	3.08 (0.80)	0.61

Note. Standard deviations in parentheses. Responses range from 1 (not at all useful) to 4 (very useful). *p* values reflect results from t-tests on the pre–post difference.

p* < .10. *p* < .05. ****p* < .01.

Looking at administrators' views from just after training to the end of the 2012–2013 school year, we found a similar pattern of beliefs (see Table 15). At the end of the year, they were least optimistic that the TGDC would support teacher dismissal decisions or increase their ability to spend more time on teaching and learning (see Table 15). They were most optimistic that it would create a common vision of excellent teaching and support teacher improvement and development. Again we tested for significant differences (at levels of $p < .01$, $p < .05$, and $p < .10$) between administrators' posttraining views of the purpose for TGDC data and their views at the end of the year, and we found that over the 2012–2013 school year, administrators' beliefs that the TGDC system could shape many human capital management decisions remained steady on four goals: teacher improvement, school improvement, teacher dismissal, and the time spent on teaching and learning. Administrators' views shifted down on three goals: the ability of the system to identify or reward strong teachers, create a common vision of teaching, and improve the district.

Table 15. Administrators' Mean Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes on Posttraining and End-of-Year Surveys

View	<i>N</i>	Post	End-of-year	<i>p</i> value
Identifying or rewarding strong teachers	212	3.22 (0.79)	3.10 (0.78)	0.04**
Teacher improvement/development	211	3.39 (0.70)	3.31 (0.71)	0.29
School improvement	209	3.33 (0.70)	3.15 (0.79)	0.08*
Teacher dismissal	205	2.95 (0.98)	2.83 (0.99)	0.13
Increasing the amount of time I can spend on teaching and learning	206	2.80 (1.00)	2.71 (1.02)	0.10
Creating a common vision of excellent teaching	211	3.46 (0.73)	3.21 (0.78)	0.00***
District improvement	210	3.14 (0.80)	2.98 (0.87)	0.01***

Note. Standard deviations are in parentheses. Responses range from 1 (not at all useful) to 4 (very useful). *p* values reflect results from t-tests on the pre–post difference.

p* < .10. *p* < .05. ****p* < .01.

Administrators' Views of the Teaching and Learning Framework

Administrators trained in the summer and school year reported broad agreement with the TLF when they were surveyed right after training (see Table A26a). They felt that the TLF generally reflected their views of teaching, covered important domains of teaching, and met their schools' specific needs. Administrators also reported feeling comfortable with the TLF. More specifically, administrators also agreed with statements such as "I feel confident in my ability to navigate MyPGS." Administrators also felt that they understood the differences between performance levels on the TLF and were confident aligning evidence. Overall, administrators felt the same when asked these questions at the end of the year, as assessed by testing for significant differences (at levels of $p < .01$, $p < .05$, and $p < .10$) between administrators' EOY views of the purpose for TGDC data and their views at the end training (see Table A29). There were no changes in administrators' views of the TLF (see Table 16) or in their views of training outcomes (see Table 17) over the 2012–2013 school year. Administrators remained supportive of the TLF at both a general and a specific level.

Table 16. Administrators' Mean Views of the Teaching and Learning Framework on Posttraining and End-of-Year Surveys

View	<i>N</i>	Post	End-of-year	<i>p</i> value
The Teaching and Learning Framework covers important domains of teaching.	153	3.80 (.40)	3.80 (.40)	0.87
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction.	154	3.69 (.48)	3.70 (.49)	0.78
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.	154	3.37 (.65)	3.45 (.68)	0.09*
I am comfortable with this observational instrument.	151	3.17 (.74)	3.14 (.83)	0.52

Note. Standard deviations are in parentheses. Responses range from 1 (disagree) to 4 (agree). *p* values reflect results from t-tests on the pre–post difference.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Table 17. Administrators' Mean Views of Training Outcomes on Posttraining and End-of-Year Surveys

View	<i>N</i>	Post	End-of-year	<i>p</i> value
I feel confident in my ability to navigate MyPGS, the online observation tool.	229	3.41 (.63)	3.00 (.82)	0.00
I understand the distinctions between the levels of performance in the Teaching and Learning Framework rubrics.	227	3.48 (.65)	3.49 (.63)	0.88
I feel confident in my ability to align evidence.	228	3.21 (.65)	3.16 (.72)	0.46

Note. Standard deviations in parentheses. Responses range from 1 (disagree) to 4 (agree). *p* values reflect results from t-tests on the pre–post difference.

* $p < .10$. ** $p < .05$. *** $p < .01$.

The one exception to this pattern was administrators' diminishing confidence in their abilities to navigate MyPGS over the 2012–2013 school year. We can speculate that training may have given administrators a somewhat false sense of confidence. During training, many people were present to support administrator learning and use of the online system. Once administrators implemented TGDC independently, they may have realized they were not as skillful as they had thought.

Table 18. Administrators' Mean Views of Their Job Efficacy and Focus on Pretraining and End-of-Year Surveys

View	<i>N</i>	Pre	End-of-year	<i>p</i> value
I can handle the daily demands of the job.	222	3.46 (.67)	3.21 (.80)	0.00***
I can focus on the quality of instruction and learning in each classroom.	219	2.98 (.83)	2.88 (.87)	0.12
I can shape the operational policies and procedures that are necessary to manage my school.	216	3.34 (.73)	3.08 (.83)	0.00***
I always use the school improvement goals to guide my decision-making.	218	3.33 (.65)	3.41 (.72)	0.04**

Note. Standard deviations are in parentheses. Responses range from 1 (disagree) to 4 (agree). *p* values reflect results from *t*-tests on the pre–post difference.

p* < .10. *p* < .05. ****p* < .01.

Administrators' Job Efficacy

New policies can create changes in the workplace environment that shape principals' efficacy and effectiveness. Therefore it is important to understand the degree to which administrators feel able to handle the demands of their jobs. Prior to training and at the end of the year, we asked administrators about their agreement with statements about the daily demands of their jobs, focusing on the quality of instruction and learning, necessary operational policies, and implementation of the school improvement plan.

Prior to training, administrators agreed or agreed somewhat that they were able to handle the daily demands of their jobs (92%), were able to shape the operational policies and procedures necessary to managing the school (87%), and always used the school improvement goals to guide their decisions (92%; see Table A15a). A smaller majority of administrators (72%) felt able to focus on the teaching and learning in every classroom. When examining results from *t*-tests comparing reports of job efficacy prior to training and at the end of the year, some views shifted over the 2012–2013 year in statistically significant ways (see Table 18). Specifically, administrators' abilities to focus on teaching and learning remained constant, but both their efficacy to handle the daily demands of their jobs and shape to operational policies and

procedures declined modestly, $p < .01$. Administrators' emphasis on using school improvement goals to make decisions increased slightly, $p < .05$.

How Did Administrators Use the Teacher Growth and Development Cycle in the 2012–2013 School Year?

Summary of Findings

Administrators used the TGDC to complete informal and formal observations. They made use of supports offered by the district, some discussed it with their colleagues, and a small number created professional development sessions for their staff (principals, teachers, etc.) about the TGDC. Despite professional development sessions offered by the education service centers (ESCs), most focus administrators felt they had insufficient opportunity and time to continue to learn the TGDC system. Most administrators (75%), however, agreed or agreed somewhat that they felt comfortable with the TGDC at the end of the year. Focus administrators implemented the TGDC as two intertwined tasks: the process of creating an observation score and the process of evaluating a teacher. Administrators' rating approaches and reasoning processes suggested they used the TGDC with both evaluation and improvement in mind.

Background

The best data to illuminate use of the TGDC comes from multiple data sources collected from the focus administrators: the EOY interview, the mid-year and EOY surveys, and the think-aloud exercise.

Interactions With the Teacher Growth and Development Cycle

Based on interview data, focus administrators had four main kinds of interactions with the TGDC. Administrators used the TGDC to complete informal and formal observations; they made use of supports offered by the district; some discussed it with their colleagues; and a small number created professional development sessions about TGDC for their staff (principals, teachers, etc.). Most administrators (79%) agreed or agreed somewhat that they felt comfortable with the TGDC at the end of the year. This left almost a quarter (21%) of administrators who did not yet feel comfortable with the TGDC at the end of the 2012–2013 school year (Table A32).

Number and Types of Observations

Focus administrators completed an average of 4.9 informal observations and 2.0 formal observation cycles during the 2012–2013 school year (Table A30). A few focus administrators ($N = 7$), however, did not work with a focus teacher at all this year. Of the 33 administrators who worked with a focus teacher, two did not conduct any formal observations. There was a range of reasons why administrators did not work with a focus teacher. Some administrators were not able to find a teacher with whom to work. Other administrators struggled with staffing changes and time constraints. These numbers are very similar to the numbers reported by administrators who returned the EOY survey.

Supports Used to Complete Observations

LAUSD offered a number of supports to administrators during the year that went beyond the weeklong training. One set of supports included e-mails and documents from the central office detailing process requirements and technical aspects of TGDC completion (e.g., making administrators aware of deadlines and requirements for the pilot year). The main support available to administrators was provided by their TLCs. TLCs are individuals who work directly with principals in their ESCs. The LAUSD central office did not mandate the content or structure of these support sessions or activities, instead leaving ESCs to decide what might be best, given the local context.

Interview data from focus administrators suggested that some districts were more successful than others in communicating with administrators and providing supports that many principals used. In some local school districts (LSDs), principals did not remember receiving any communications from the LSD and did not seem to know that professional development sessions were being offered. In others, they did know about support sessions but chose not to attend. Most focus administrators knew that they could call or e-mail their TLC, and those who did call or e-mail felt strongly supported. More than half of focus administrators (56%) did not use or modestly used individuals who were designated to support them (see Table 19). Specifically, there was no or little use of instructional directors (55%), TLCs (62%), and central office staff (69%).

Table 19. Extent of Resource Use to Support Teacher Growth and Development Cycle Implementation, 2012–2013

Resource	Not at all	A little	Some	A lot
My instructional director	44.7% (17)	10.5% (4)	36.8% (14)	7.9% (3)
My teaching and learning coordinator	30.8% (12)	30.8% (12)	23.1% (9)	15.4% (6)
Central administration staff	56.4% (22)	12.8% (5)	25.6% (10)	5.1% (2)

Note. $N = 38$. Frequencies are in parentheses.

These survey findings were supported by interviews with the focus administrators. When asked what kinds of supports they needed moving forward, with one notable exception, focus administrators said the district provided insufficient opportunities to continue to learn to do this evaluation work or that they were unable to avail themselves of the opportunities that were offered. A number of administrators suggested that opportunities to watch videos of lessons and talk with one another about ratings and/or coaching teachers would be particularly helpful to them. Dafne's, Jason's, and Anthony's (pseudonyms) responses suggest that at least some principals are hungry for opportunities to learn from one another.

I would love the opportunity to talk to other administrators to talk about instruction. It was very apparent that this had never been used system wide in any kind of system. Talking about instruction would be great. (Dafne)

Analyzing, meeting, finding out what administrators are most comfortable with [are the things I need to support my learning]. Right now, it's kind of top-down. I would want to get feedback, honest feedback. (Jason)

We are sorted out in networks but we don't get together and do things together. I don't know of any of the 12 others [principals] who get together. We meet monthly in networks and we never network about anything. The meetings are usually professional development mainly in Common Core and the EL plan and you just go back to your site. There was a lady who is the math coordinator and I asked if she wanted to do an exchange with me. More networking with the expertise of different principals would help. (Anthony)

Another resource that was in short supply was time. Principals in particular asked for more time to learn the system. In interviews, they described the reality of busy days filled with both operational and instructional demands. This led to not enough time to learn the TGDC. Rose emphasized this need for time: "You have to give the principals opportunities to learn how

to really use the system.” Kyle emphasized the need for places to practice that would be safe and provide principals feedback, in his words, “sandboxes.” When asked what he needed to learn, Kyle said, “A sandbox that’s accessible, readily available, where administrators can do practice. It would allow you to measure your own growth and allow people to be safe learning.”

Use of the Teaching and Learning Framework During Observations

We were not able to co-observe with administrators when they were actually conducting preobservation conferences, classroom visits, and postobservation conferences, so our understanding of how administrators used the TLF is limited to survey, interview, focus group, and think-aloud data. For full descriptions of the methods and instruments, please see the section, *How Was This Report Prepared?*, and Appendix C.

As mentioned earlier, administrators’ use of the TLF can be divided into two intertwined tasks: the process of creating an observation score and the process of evaluating a teacher. It is important to note that most administrators did not see these as two separable tasks. And as described subsequently, the two tasks overlapped and informed one another for many, although not all, administrators.

Score Creation Process

As described previously, training did not require administrators to follow a specified process to determine their ratings. Trainer modeling about how to rate teaching practice consisted of the trainer going back and forth between the evidence from the lesson (the transcript and the video) and the language of the TLF rating scales (i.e., the rubrics). Administrators created their ratings in one of at least two ways: working back and forth between the rating scales and the evidence and relying on their own understanding of the rating scales and aligning the evidence they took with the rating scale descriptions.

In the first approach to rating, administrators went back and forth between the words in the TLF (e.g., the words used to describe “basic” performance) and the evidence observed during the lesson. An example of this back-and-forth strategy from Heather, an elementary school principal, is provided in the following. While rating the Quality and Purpose of Questions element, she began by reading the definition. Then she read the evidence from the transcript. In the third paragraph, Heather went back to the rubric and considered whether she had evidence to support the rating. She said:

[Reading from the definition of the element:] Cognitive challenge, all students participate. Embedded in the lesson is that students are posing questions. I highlight all of these things and student–student discussions. So now I go back into the transcript and look for examples of each one of these things.

OK, cognitive challenge [reading through the transcript looking for evidence. She finds evidence at 1:15 and then the text at 5:24.] I like this one quite a lot. He is posing them questions rather than just telling them facts. Such as at 9:09 and 10:00. [She reads evidence at 9:09 and 10:00 out loud.] Specifically, I like “But tell me why.” I like the positive challenge and that all students are participating.

[Looking at the rubric again] I didn’t see them [the students] asking questions. So I would say no, that didn’t really happen much. Student–student discussions. I don’t see too much of that. If the teacher is asking questions and students are answering, they’re just facilitating. I don’t see students talking to each other. I didn’t hear him even say that they should be asking each other questions. I have no evidence for that.

In the second approach to rating, administrators referenced the language of the TLF rarely, if at all, instead relying on personal criteria for rating the element. They did not reference the rating scales, but they did have reasons for the ratings they gave. Those reasons were often related to the rating scale definitions. Harvey justified his reasoning by saying:

I saw a little differentiation with different groups, but not as much as I would have liked to see, in the sense of . . . there were not any clarifying questions. He was just going through our process of thinking, but didn’t go a lot more into depth. If a kid was struggling, he wouldn’t have gotten it. These are kids who got it, what about kids who didn’t? It wasn’t really clear, just in taking a snapshot, I didn’t see anything to support students who are still struggling with that issue.

When Harvey was rating the Quality and Purpose of Questions element, he did not work back and forth between the rubric and the evidence. He commented about how much differentiation he would like to see but did not connect this to the scoring criteria. Instead, he specified the behavior he considered evidence of differentiation: clarifying questions. Harvey went on to explain the outcome of not asking clarifying questions.

In addition to these more general strategies, many administrators began their scoring work by reading the *effective* category of the element they were working on. Administrators reasoned that “effective” is “where you want teachers to be,” so one should begin there. We do not know whether this approach led to more accurate and valid ratings, but it was widely used.

While some administrators reasoned to ratings using evidence and the rubric, the process was not automatic. Most of the think-aloud interviews were conducted between March and June 2013, and many administrators reported that they had not used the TLF since training 8 months prior. While the informal observations were supposed to be conducted with the TLF, some administrators reported they did not do so, instead using their prior evaluation practices. Thus it seems likely that the formal observations constituted the majority of administrators’ practice with the TLF during the 2012–2013 school year. Further analyses are not possible because administrators were not required to document informal observations in the same way they documented formal ones in MyPGS.

Administrators also mentioned and showed uncertainty with specific core concepts of the TLF. The think-aloud protocol only asked administrators to work with three elements—Expectations for Learning and Achievement, Quality and Purpose of Questions, and Standards-Based Projects, Activities, and Assignments. Within these elements, administrators were uncertain about differentiation and culturally relevant pedagogy—two concepts that cut across elements. These concepts were the most frequently nominated as needing greater clarification:

[Reading from the rubric:] The task is culturally relevant. I don’t know what that means. Are they talking about culture in the ethnic sense or the culture of academia? (Anthony)

We don’t have a lot of common language for what culturally relevant pedagogy, for example, should look like. If we are trying to create a norm and a standardized way of doing this, we really need to invest in having a common scoring standard. We make too many assumptions about these things based on our own preconceived notions, and it’s very challenging without that common language. (Heather)

Finally, most administrators were concerned about the directive to script lessons. While they acknowledged that scripting produced invaluable evidence for productive feedback sessions with teachers, many also felt that scripting was too intensive and made them miss important aspects of classroom interactions. One principal explained that she felt she missed her regular

check-ins with students by being so focused on scripting. Her pre-TGDC practice of walking around the room, listening to students, and asking students to explain why they were doing what they were doing was replaced by scripting. She explained that she even walked around the room with her laptop, trying to capture every word the teacher said, as she visited students. More than one principal mentioned that, when focused in this way on verbal interactions, he or she missed the nonverbal interactions and evidence (e.g., looks of confusion or understanding, what was written on the board, physical interactions between teachers and students). Some also felt it was difficult to track what the lesson was about when they were busy scripting. Some of these administrators felt that this losing track of important pieces of information made the job of providing accurate ratings and high-quality feedback more difficult. Many administrators were optimistic that their scripting would improve with time, but they also found it challenging. Portia's explanation was representative of others' concerns:

I cannot possibly get everything that the teacher is saying and I don't know how well I'll do putting it onto the platform, because I might not be able to have the required evidence to adequately complete it. For that I feel bad.

Evaluating a Teacher

Administrators had many tasks related to observing. One was to schedule classroom visits and observe lessons, scripting teacher and student talk verbatim. Another was to create the ratings. Yet another was to have conversations about those ratings with teachers for the improvement of practice. One might reasonably think that these were separate tasks—an administrator created a rating and then had a conversation about it. The think-aloud data, however, suggested that administrators carried out these tasks in a more integrated way. This finding stood in contrast to research studies that have used researchers, not principals, to create ratings (e.g., Bell et al., 2013).

When interviewing focus administrators and watching them use the TLF during the think-aloud interviews, we found they were often thinking about observing with the purpose of evaluation in mind. They were not only thinking about creating ratings but also thinking regularly about how they were going to have the postobservation feedback session or ongoing professional conversations across the school year. They also thought about how they might help teachers improve their practice while they were observing and rating.

Specifically, when administrators were asked to explain why they gave the rating they did, they often referred to strategies such as the ones described previously, but they also often made reference to the outcomes of the ratings, for example, how they envisioned the conversation with the teacher might go, what the administrator might say to the teacher, or what the administrator's general approach to the rating conversation would be. For example, when Sara was describing how she would discuss the ratings she gave to the teacher, she said, "If this had been my teacher, I would have had him looking at the transcript. And we'd go down everything that was said. I'd tag with the teacher. It takes a long time."

Anthony told a story about how he had worked with a teacher whom everyone in the building perceived as "highly effective," which, in part, was why he asked her to participate in the pilot year. He found over the course of the year, however, that some of her practices were not exemplary and that there were some areas on which she could improve. He explained that, after the first observation in which he was honest with this teacher about the areas where she might grow, the teacher stopped talking to him and did not want to participate in the pilot any longer. Anthony explained that, in this experience, he had to think of the "human drama" a particular rating would create. He said:

When I gave her *developing*, she was not happy. I don't worry about that. For the most essential stuff, I gave her *developing*. She was far away from being *effective*. I think she's a teacher that thought of herself as *highly effective*. That is important because I have to think about the human drama that this is going to create. [Italics added to note rubric choices.]

Administrators were very aware of and discussed the fact that evaluation was a human endeavor. They agreed that they had to consider politics and personalities when conducting observations.

Some individuals developed rating processes to help them have productive evaluation conversations with teachers. For example, Heather's process of taking evidence and coding was designed around how she would share the evidence and ratings with her teachers. She explained the process she used and noted that her coding work helped teachers see what she did. She said,

I may not have time to score it right away, so I go through with a highlighter. I create a little legend and I will highlight 2b—I'll use a blue highlighter. Yellow may be 3c. I'll go through all of my colors on the hard copy. That's just a backup for me so that I don't miss

anything. If I get confused or if I try to figure out where the part was, I have a highlighted piece. It's my safety net and the highlighting makes it fresh in my mind. And it's also something I show the teachers. I show them the highlighted version. I want those to be conversation pieces. I want them to see all of the yellows.

Another common strategy administrators used was to provide more evidence than they believed strictly necessary. They explained that, by showing the teacher that the principal had been paying attention and taking careful notes, they hoped to decrease the likelihood the teacher would feel the principal was not being objective and fair in the evaluation.

In addition to anticipating the postobservation conference or evaluation context, administrators sometimes noted what the teacher should have done differently. This was most common when the administrator was asked to justify or explain why he or she gave a certain rating and occurred occasionally when the administrator was actually watching a video of practice. For example, in describing the questioning technique the teacher used, an administrator might note how the teacher could have gotten more students involved or how the teacher could have rephrased the question so that it was more cognitively challenging. Examples from administrators included the following:

It would have been good if he got students to be clearer and got it [the answer] out of students instead of giving it to them. (Sara)

In order to prove a little bit of my point, I'd cut and paste 5:24 [referring to the evidence in the transcript at time marker 5:24] and I'd say, "Here would have been the perfect opportunity to ask students to have them paraphrase to make sure they knew what they were supposed to do." Even the strategy of breaking up into groups—at the last minute sorting kids into groups—in order to be effective, those kids would get into groups in a more organized and instant way. It wasn't that bad, but manipulatives could have been organized ahead of time. (Anthony)

At least in the 10 minutes, he did not have any classroom management problems. He had a few systems in place that could have been better. In terms of picking up materials. And I saw that he lacked in terms of seeing that everyone had the materials. He should have had an overhead where everyone could read or be able to see. There were just a few things lacking. (Ella)

This practice of noting what the teacher should have done to justify a rating was one we saw rarely in previous think-aloud work with raters who were researchers rather than local administrators (Bell *et al.*, 2013). Though we can only speculate, it is possible that LAUSD administrators thought about what the teacher could have done differently in part because the administrator was responsible for helping improve instruction.

Summary Findings and Recommendations

Administrators and focus administrators shared many insights throughout the 2012–2013 practice year. These insights are provided as follows in a bulleted, high-level summary, followed by recommendations.

Training

- Administrators came to initial training with experience conducting informal observations to help teachers improve. They were experienced at identifying strengths and weaknesses in teaching practice. They had less experience coplanning lessons, reviewing lesson plans, or suggesting readings or curriculum resources to support teacher learning.
- After the initial training, 81% of administrators reported that they learned a lot about how observations fit into the TGDC.
- Many administrators said they learned some or a lot about how to take better notes (81%), accurately score observations (96%), use MyPGS (90%), and give feedback to teachers following observations (84%).
- Administrators reported learning less in areas not directly covered in training. Between one half and two thirds of administrators reported learning little or nothing about how to manage the TGDC with their other responsibilities, how to match teachers with professional development opportunities, how to effectively manage the strengths and weaknesses of their staff, and how to help all people learn to improve.
- During training, administrators were unsure about a number of issues that are important to the observation scores resulting in changes in teaching practices, for example, what professional development will be available for teachers when needs are identified through the observation system, how to have conversations with various

groups of teachers, how many teachers will need to be observed yearly, or how to incorporate these new requirements into their already-full schedules.

- Most administrators were very satisfied with training, and there were no systematic differences across grade levels, job roles (central office staff, principal, etc.), or when the administrator was trained.
- Administrators who believed more in the importance of teacher evaluation and had higher incoming expectations for training reported higher levels of satisfaction with training. Administrators who said their jobs were more manageable reported somewhat lower levels of satisfaction.
- Between Summer 2012 and Summer 2013, several changes were made to the TGDC trainings. They began to include more conversations about the differences among performance levels, a full day of training on how to have coaching conversations with teachers, and more emphasis on TLF language. Trainings also resulted in more posttraining confidence with MyPGS over time.

Certification

- Across the two certification videos, 7% of administrators were certified after the initial training. The majority of administrators were preliminarily certified (~82%), and 11% of administrators were not yet certified after training.
- Of the individual components of certification, administrators were most likely to become certified on objectivity. Accuracy appeared to be the hardest component for administrators, with nearly one third not yet certified.
- Performance on the accuracy component of certification revealed that different teaching acts could be variously interpreted, and administrators do not hold views that are consistent with those of master raters or of one another.
- Administrators agreed with one another and master raters most on Standard 2, which is principally concerned with behavior management, organization, and the classroom environment.

- Administrators agreed least on elements in Standard 1 (Planning) and Standard 5 (Professional Growth) rather than Standard 3 (Standards-Based Instruction). On these three standards, administrators tended to score the teacher higher than master raters. On Standard 2 (Designing Coherent Instruction), administrators scored slightly lower than master raters.
- Principals had higher agreement rates than assistant principals; elementary and secondary principals had similar agreement rates.
- Though principals were most likely to become certified as compared with other role groups, the gap between various groups' certification performance was not large. The same is true of school level: Elementary principals are more likely to become certified than secondary principals, but the gap is small.
- Aside from one's view of teacher evaluation, administrators' pretraining beliefs were not related to their likelihood of overall certification. Administrators who valued teacher evaluation coming into training were slightly less likely to become certified than those who reported less belief in the importance of teacher evaluation.

Conducting Observations and Using the Teacher Growth and Development Cycle

- Administrators have a range of knowledge and skills that must be accounted for in the implementation of TGDC. Specifically, they have relatively little experience working with teachers in detailed aspects of improving their practice and have more experience making higher level judgments about the quality of teaching.
- Administrators spend considerably less time on tasks related to work with teachers or activities directly related to teaching than they spend on tasks such as operations and paperwork. For some administrators, this means they are not spending time on the things they view as most critical to helping teachers improve.
- Administrators generally feel able to handle the demands of their jobs, although they generally report lower ability to focus on and shape instructional tasks as compared to operational tasks.

- Administrators generally viewed the TLF itself and the TGDC practice year experiences very favorably. The TLF was aligned to their views of quality teaching. It helped them think more critically about teaching, and they believed it will be helpful in supporting teacher development.
- Most administrators (75%) felt training and the yearlong practice left them prepared for the 2013–2014 implementation.
- Roughly a quarter of administrators (26%) who returned the EOY survey did not work with a focus teacher in 2012–2013. Most administrators (68%) did conduct at least one formal observation with a teacher. A few administrators (12%) said it was difficult to find a teacher with whom to work.
- There was wide variation in the number of formal and informal observations conducted as well as how long it took to complete them. On average, administrators who returned the EOY survey conducted two formal and five informal observations with their teachers. Formal observations took 10 hours each on average, and focus administrators reported that informal observations took 2 hours each.
- Most administrators (75%) reported that it took too much time to conduct a formal observation.
- Focus administrators most often requested additional support for learning how to align evidence to specific TLF elements and to manage the transition to full TGDC implementation.
- Administrators raised two major concerns about conducting observations: (a) the time to conduct formal observations and (b) confidence on the technical aspects of the MyPGS platform.
- Administrators are most optimistic that the TGDC can support teacher improvement and development. They are least optimistic that it will support teacher dismissal decisions or increase their ability to spend more time on teaching and learning.
- Focus administrators had four types of TGDC experiences: completing informal and formal observations, making use of various resources to learn the TGDC, discussing

the TGDC with colleagues, and incorporating the TGDC in professional development for their teachers or staff.

- ESC support for continued learning was uneven. There were ESCs in which administrators felt there were opportunities to continue to learn; however, most focus administrators reported they had insufficient opportunities and time to continue to learn the TGDC. Almost a quarter (21%) were not yet comfortable with the TGDC moving into the 2013–2014 school year.
- Focus administrators' work with the TGDC can be divided into two intertwined tasks: the process of creating an observation score and the process of evaluating a teacher. These two tasks overlapped but were not the same.
- Focus administrators used the TLF with the purpose of evaluation in mind as well as how they might help a teacher improve that teacher's practice. Administrators did not simply create a score with no regard for its intended purpose.
- Focus administrators viewed evaluation as a human endeavor and reported they had to consider politics and personalities when conducting observations.
- Focus administrators asked repeatedly for time and opportunity to work with other administrators to learn how to use the TGDC.

Changes in Administrators' Views

- In general, administrators' views did not change dramatically over the year of training in and practice with the TGDC.
- Specifically, administrators' views of the TLF did not change.
- Administrators' beliefs about how well the TGDC could contribute to the district's many human capital management decisions remained steady on four goals (teacher improvement/development, school improvement, teacher dismissal, and public accountability) and shifted down on three: the ability of the system to identify or reward strong teachers, create a common vision of teaching, and improve the district.

- Administrators became less confident in their abilities to navigate MyPGS, manage their jobs, and control operational decisions. Except for the decrease in confidence about using MyPGS, these reductions were small.

Recommendations Summary

Training

- Continue to build on high-quality and effective aspects of training while refining training based on pre- and postsurvey and certification results.
- Directly address administrators' concerns about the manageability of the TGDC with their existing responsibilities. This might include specialized staff to conduct observations for certain groups of teachers, the use of peers in conducting informal observations, and combining formal and informal observations in ways that reduce time demands but still support teacher and administrator learning.
- Provide specification about how to match high-quality professional development opportunities to teachers' areas of need.
- Continue to provide hands-on MyPGS support.
- Continue the weeklong initial training for new administrators as well as the certification test.

Certification

- Given the history of near-universal proficiency in most teacher evaluation systems across the country, focus resources on offering additional professional development sessions for administrators who are NYC and PC on the accuracy component. To the greatest degree possible, use administrators who are *certified* and *certified with distinction* to support the learning of others.
- Continue to offer opportunities for administrators to learn more and try to become certified again in other components of the system.
- Continue to work on developing both a shared language for teaching and a shared sense of what teacher and student behaviors demonstrate proficiency on the TLF.

Conducting Observations and Using the Teacher Growth and Development Cycle

- Administrators believed the TGDC will support teacher development and they were eager to use it for this purpose. Build on this momentum by providing strong professional development for teachers aligned with their areas for growth. Provide strong professional development for administrators about having conversations with teachers about how to improve their practice via the TGDC.
- Provide better communication and more centralized tracking of the substance of ESC-led TGDC trainings. Track how many administrators take advantage of these trainings and their views of the trainings after having attended.
- Create opportunities for administrators to score common lessons and compare their appraisals of these lessons to hone their skills in interpreting teaching acts, distinguishing among levels of teaching proficiency, and build shared understandings of high-quality instruction.
- Widely and aggressively publicize learning opportunities for administrators to improve their comfort and skill with TGDC. Professional development opportunities should also be varied because administrators have different learning needs.
- Investigate why administrators believe TGDC will be unlikely to support teacher dismissal decisions.
- Acknowledge and provide support for the range of teachers administrators will support via the TGDC. Administrators see teachers as individuals, and administrators need help learning how to address differences productively.
- Pay careful attention to administrators' concerns about their ability to manage their jobs and widespread agreement that the TGDC took up too much of their time. LAUSD enjoyed widespread and strong support for the substance of the TGDC. Administrators saw it as a tremendous resource that should not be squandered. Develop appropriate responses to these concerns that do not compromise implementation quality.

How Was This Report Prepared?

This report is based on data collected from a number of sources from July 2012 to June 2013. Five training sessions were observed, with extensive field notes recorded; some were video-recorded. All administrators were surveyed using online questionnaires before, during, and after the summer training. A group of 42 focus administrators filled out more extensive online questionnaires and participated in interviews and think-aloud exercises. Here we describe each of these instruments. The instruments themselves are included in Appendix C of this report.

Training Surveys for All Administrators

During the 2012–2013 school year, we administered surveys to all training participants before, during, and after their training. The enrollment list provided by LAUSD before each training session was used to determine who would get the survey. The surveys were all administered online through SurveyMonkey. After the initial invitation, a reminder was sent to those who did not respond within 2 days. Some of the administrators who attended two trainings and thus showed up in multiple enrollment lists only needed to fill out the surveys once. Administrators were provided with an opt-out option that allowed them to block any future training survey invitations. At the end of the school year, we also administered an EOY survey to all training participants on behalf of LAUSD.

We developed the surveys with input from administrators, educational leadership faculty, and experts in the area of teacher observation, in conjunction with TLS. Each instrument was piloted with current non-LAUSD administrators to estimate the length of time required to complete each survey and to identify whether questions would benefit from further revision.

Pretraining survey: All administrators. All administrators attending trainings received an e-mail invitation to complete one online survey 2 days before their training session. The pretraining survey was composed of 12 questions. Its primary goal was to collect information ahead of training that would affect administrators' engagement in and success with the training content. During the 2012–2013 school year, the pretraining survey was sent out to 820 administrators who went through the training. The response rate for the pretraining survey was 82.6%. This instrument is found in Appendix C.

Mid-training survey: All administrators. On the third day of each training session, usually Wednesday of the training week, administrators received another e-mail invitation to complete an online survey. This is also referred to as the Day 3 survey, which included 13

questions. It was administered on behalf of TLS and served two purposes: to assess administrators' learning and to assess their satisfaction with the first 3 days of the training. There were 947 Day 3 surveys sent to administrators who went through training. The response rate for the Day 3 survey was 41.1%; this response rate was lower than the rates for the pre- and postsurveys, likely because the window during which participants could respond was truncated. This instrument is found in Appendix C.

Posttraining survey: All administrators. On Monday after the training week, administrators received another e-mail invitation to complete the posttraining survey online. The posttraining survey had eight questions and focused on assessing administrators' (a) learning, (b) confidence, and (c) satisfaction in the training. There were 966 posttraining surveys administered to people who went through the training. The response rate for the posttraining survey was 65.5%. This instrument is found in Appendix C.

End-of-year survey: All administrators. The EOY survey was sent out to all administrators on behalf of LAUSD on June 19, 2013, via SurveyMonkey. The survey was developed collectively by the research team with input from the district and included nine questions. It mainly focused on administrators' self-assessment of understanding and using the TGDC process, their experiences conducting formal and informal observations in 2012–2013, their impression of the TLF, their opinions of how the observation component of TGDC might be useful, and their perceptions of how they can and will be able to cope with the various aspects of their job. Out of 1,182 administrators, 397 (34%) responded to the survey (Table 20); the response rate, although low, is in line with typical response rates when instruments are not mandatory (e.g., Rockoff, Staiger, Kane, & Taylor, 2012). This instrument is found in Appendix C.

Table 20. Survey Schedule and Response Rates

Survey	When delivered	No. questions	Response rate (%)		
			Overall	Summer	School year
Pretraining	Friday before the training week	12	82.6	81.8	84.6
Day 3	Wednesday of the training week	13	41.1	41.4	39.9
Posttraining	Monday after the training week	8	65.5	66.8	60.4
End-of-year	June 19, 2013	10	33.9	NA	NA

Surveys, Interviews, and Other Data for Focus Administrators

In addition to the surveys for all administrators, we collected more intensive survey and interview data from a sample of 42 focus administrators. These included an initial interview

during training, an expanded posttraining questionnaire, a think-aloud interview session, three bimonthly short surveys, and a final spring interview. Of the 42 focus observers, we have complete data from 38.

We developed the instruments with input from administrators, educational leadership faculty, and experts in the area of teacher observation. As with the other instruments for this study, each instrument was piloted with current administrators to estimate the amount of time required to complete each survey or interview and to identify whether any questions would benefit from further revision.

Training questionnaire: Focus administrators. After the weeklong training was complete, all focus administrators received invitations to complete an online survey with questions in the following categories: (a) administrators' perceptions of their skills and abilities in conducting teacher evaluations/observations; (b) administrators' practices when evaluating classroom instruction prior to the TGDC; (c) administrators' perceptions about the TGDC system and the TLF; (d) administrators' experiences working with teachers as an administrator; (e) administrators' opinions about administrator actions in supporting teachers; (f) time spent at work, including conducting teacher evaluations/observations; and (g) satisfaction level with job and life. This instrument is found in Appendix C.

The survey was administered online using ETS's proprietary survey platform. Each focus administrator received an e-mail with a link to the survey. The e-mail request for filling out the survey was sent out on July 12, 2012, and two reminders were sent out in late July and early August. We received 41 survey responses, with two people filling out the survey twice. Therefore we have 39 responses out of 42 focus administrators.

Training interview: Focus administrators. Forty-one of the 42 focus administrators were interviewed by research team members, and the interviews were audio-recorded with interviewees' consent. The interviews took around 30 minutes, and the majority of them were conducted on-site during the training. A few were conducted on the phone shortly after the training.

The focus administrator training interview protocol was developed collectively by the research team and included seven open-ended interview questions. It mainly focused on administrators' job responsibilities, their past experiences evaluating teachers, their views on

good teaching, and their expectations on LAUSD's TGDC system as well as its implementation. The interview questions are found in Appendix C.

Bimonthly surveys: Focus administrators. The first bimonthly focus administrator survey was sent out to all 42 focus administrators on February 13, 2013, via SurveyMonkey. The survey was developed collectively by the research team and included five short questions. It mainly focused on checking administrators' progress and experiences with conducting formal or informal TGDC observations thus far and their opinions on how the observation training had prepared them, and it also polled administrators for what would be a good time to visit them in Los Angeles for in-person interviews. Thirty-nine out of 42 focus administrators responded to the survey. This instrument is found in Appendix C.

The second bimonthly focus administrator survey was sent out to all focus administrators on April 17, 2013, via SurveyMonkey. The survey was developed collectively by the research team and included four short questions. The main focus of this survey was on checking administrators' progress and experiences with conducting informal TGDC observations thus far, their opinions on how they view informal observations in their jobs, and the areas they most need training on in the 2013–2014 school year. Thirty-eight out of 42 focus administrators responded to the survey. This instrument is found in Appendix C.

The third bimonthly focus administrator survey was sent out to all focus administrators on June 12, 2013, also via SurveyMonkey. The survey was developed collectively by the research team and included four short questions. It mainly focused on understanding administrators' expectations for the next year in terms of how important and time consuming the various aspects of their job will be, their opinions on what types of learners the teachers/principals they supervise are, and to what extent they draw on various individuals or resources to learn about the TGDC. Thirty-nine out of 42 focus administrators responded to the survey. This instrument is found in Appendix C.

Think-aloud: Focus administrators. Think-aloud sessions were conducted with 35 out of 42 focus administrators in Spring 2013. They took approximately 1 hour, and most sessions were conducted on the phone with the aid of GoToMeeting, which enabled researchers to observe administrators as they aligned evidence and scored. A few sessions were conducted in person when research team members visited Los Angeles.

The think-aloud protocol was developed collectively by the research team. Focus administrators were asked to watch a 10-minute video of classroom instruction, with transcripts of the video provided to them. Then they were requested to align one piece of evidence to each of the three focus elements we selected and assign scores using the TLF rubric as they normally would do in an observation cycle. In this process, they were asked to think out loud as much as possible. When they completed scoring, a stimulated recall session was conducted in which researchers asked specific follow-up questions about how the focus administrator was thinking about specific scales (e.g., why a certain score could not be higher or lower than what the administrator assigned; how the administrator used and evaluated the evidence collected). At the end of the session, we also asked them to provide feedback on the process of rating as well as on experiences using the rubric. This instrument is found in Appendix C.

Spring interview: Focus administrators. Twenty-six out of the 42 focus administrator were interviewed by research team members in Spring 2013, and the interviews were audio-recorded with interviewees' consent. The interviews lasted between 30 and 60 minutes, and the majority of them were conducted by phone. A few were conducted in person when research team members visited Los Angeles. The spring interviews were designed to be conducted once principals had conducted at least one observational cycle with at least one teacher. For some administrators, this was not the case, because they had not secured a teacher with whom to work.

The focus administrator spring interview protocol was developed collectively by the research team and included 13 open-ended interview questions. It mainly focused on administrators' goals of conducting TGDC observations, their relationships with teachers and how the TGDC impacts those relationships, their experiences learning the TGDC, their needs in learning and using the TGDC, and their perceptions of the evaluation policy. The interview questions are found in Appendix C.

Other Data Sources

In addition to data from all administrators and from the 42 focus administrators, we collected data from central office staff and focus groups of principals and instructional directors.

Interviews of central office personnel. Six key individuals from the LAUSD central office were interviewed in late Spring 2013. The central office personnel were selected for their close involvement with the TGDC implementation. Questions were tailored to the particular roles the staff members played and were generated collaboratively by the research team based on

weekly meetings. Questions covered a wide range of topics but were designed to gather the unique perspective that staff members had on implementation. All interviews were used as background conversations to inform our understanding of the TGDC and emerging findings.

Focus groups. At the request of central office personnel, two focus groups were held in April 2013, bringing together instructional directors in one group and principals in another. The central office consulted various stakeholders to solicit names, and all names that were provided to the research team were invited to participate. Two TLCs were at both focus groups. Five principals attended their focus groups, and seven instructional directors attended their focus groups. The focus group sessions were recorded and transcribed.

Data Analysis

Many analyses were conducted in the first year of the study. Some descriptions of analysis procedures appear in previous sections of this report. More detailed descriptions follow.

Survey analyses. Survey measures were developed based on existing survey measures and drew on guidance from educational leadership faculty members and principals currently working in the field. For all of the surveys administered, data were first cleaned to identify outliers and patterns indicating that data were incorrectly coded. Next, a codebook was created summarizing means and standard deviations for all numeric variables and tabulations for all categorical and dichotomous variables. To reduce the number of survey items included in our analyses, we calculated mean composite scores from multiple related survey questions; these mean composites were qualitatively similar to results from factor analysis, the goal of which is to identify latent constructs underlying categories of variables. In all survey analyses, we included several control variables to account for characteristics of raters (e.g., instructional level) and training (e.g., week, location) that could potentially affect the association between key survey variables of interest and our outcome variables, namely, satisfaction with training and certification. We also used robust standard errors to better account for the shared experiences of administrators within the same training site.

Think-aloud analyses. Think-aloud interviews were discussed regularly at weekly project meetings prior to formal analyses. Members of the research team took as close to verbatim transcripts as they could during the think-aloud interviews. First, these notes were read holistically and the research team developed a set of emergent themes. Those themes included the ways that the administrator arrived at a score (e.g., did they use the TLF, did they use their

personal expertise, how did they use evidence), ideas about how to improve teaching, how administrators justify a score, understanding of core concepts, degree of comfort and automaticity with the TLF, and how the scores would be used. Next, 13 think-aloud interviews were coded based on each of the preceding themes. There were additional themes we did not have time to go back and code for (e.g., accuracy of evidence use), and those themes will be further analyzed at a later time. An analysis memo was written and reviewed. This memo formed the basis for the analysis included in this report.

Interview analyses. Interviewers took verbatim notes during all interviews; these notes were checked against the recorded interviews for accuracy. Training interviews were coded for main themes. Those themes were summarized, and a training interview summary was written for the project team. A sample of 13 of the 42 was analyzed using qualitative data analysis software Dedoose for the purposes of this report. The sample of 13 was picked from among all 42 interviews; the sample was selected so that interviews from all interviewers would be represented, as well as a range of participants with different job titles, including principals, assistant principals, and coaches from elementary, middle, and high school. This sampling was done due to time constraints. A full coding of the interviews will be conducted for subsequent analyses. A summary memo, using quotations to illustrate themes, was created for the study team to discuss. This memo formed the basis for the analysis included in this report. Because each member of the study team conducted these interviews and reviewed the spring interview memo, we have no reason to believe the random sample misrepresents the main themes in the interviews.

Writing

This report was collaboratively written by the UCAST research team. Team members worked to triangulate data from different data sources, coordinating quantitative and qualitative data analysis and presentation as much as possible.

Glossary

To assist the reader, we define terms specific to LAUSD or to our study that are used throughout. A more detailed description of each of the terms in this glossary is included in the body of the report.

Evaluation Context

certification. All administrators in the district were required to be certified on the TLF. The TGDC training culminates in a certification test in which administrators review a lesson plan, watch the video-recorded lesson, and review the postobservation conference.

Education Service Center (ESC). LAUSD has four sites for training and other educational services spread out geographically (north, south, east, and west) where summer and school-year TGDC trainings were offered.

Educator Growth and Development Cycle (EGDC). The EGDC is the overarching evaluation system in LAUSD, with specific versions tailored to evaluate central office staff, school-level administrators, and teachers.

focus elements. To streamline the work of scoring observations, LAUSD identified a subset of elements to score on the TLF, and these were referred to as “focus elements.” In the 2012–2013 school year, LAUSD included 21 focus elements in the TLF.

local school district (LSD). LAUSD is divided into eight smaller district organizations, or LSDs, which are determined by geographical zones of family residences.

My Professional Growth System (MyPGS). MyPGS is the online platform into which administrators entered data (e.g., scripted evidence from observations, final observation ratings) for the TGDC evaluation process.

School Leader Growth and Development Cycle. For school leaders, including principals and assistant principals, this process is parallel to the TGDC process. It consists of formal observations to assess the quality of administrators’ and district officials’ pre- and postobservation conferences with teachers.

STULL. This is the observation-based evaluation system that predated TGDC. In the STULL process, areas of evaluation included achievement of instructional objectives, preparation and planning, classroom performance, and general professional skills. STULL ratings were binary: meets expectations/does not meet expectations.

Teacher Growth and Development Cycle (TGDC). The TGDC is the teacher evaluation system we refer to throughout the report. The TGDC consists of formal goal setting at the beginning of the year, two formal observations with pre- and postobservation conferences, informal observations, and a formal summary evaluation.

Teaching and Learning Framework (TLF). The TLF is LAUSD's name for its observation protocol, which is a modified version of Danielson's Framework for Teaching (Danielson & McGreal, 2000).

TGDC training. A mandatory 4-day training exercise was completed by all administrators prior to conducting TGDC observations. Most participants in our study were trained during the summer, although a small number were trained during the school year.

Personnel

administrator. An administrator is any school- or district-level leader tasked with conducting formal TGDC observations. The administrators in our study are primarily school leaders (i.e., principals and assistant principals).

focus administrator. Focus administrators are the 42 volunteer administrators who were selected to participate in additional data collection activities for the research team, including surveys, interviews, and think-aloud exercises.

instructional director (ID). Each LSD has an instructional director who coordinates evaluation activities and supplemental training for administrators within that district.

teaching and learning coordinator (TLC). TLCs are individuals who were hired to support administrators in their local school districts.

References

- Bell, C. A., Qi, Y., Croft, A. J., Leusner, D., Gitomer, D. H., McCaffrey, D. F., & Pianta, R. C. (2014). Improving observational score quality: Challenges in observer thinking. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 50–97). San Francisco, CA: Jossey-Bass.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Jerald, C. (2012). *Ensuring accurate feedback from observations: Perspectives on practice*. Seattle, WA: Bill and Melinda Gates Foundation. Retrieved from <https://docs.gatesfoundation.org/Documents/ensuring-accuracy-wp.pdf>
- Kee, K., Anderson, K., Dearing, V., Harris, E., & Shuster, F. (Eds.). (2010). *Results coaching: The new essential for school leaders*. Thousand Oaks, CA: Corwin Press.
- Rockoff, J. E., Staiger, D. O., Kane, T. J., & Taylor, E. S. (2012). Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review*, 102, 3184–3213.
- Sartain, L., Stoelinga, S. R., Brown, E. R., Matsko, K. K., Miller, F. K., Durwood, C. E., . . . Glazer, D. (2011). *Rethinking teacher evaluation in Chicago*. Chicago, IL: Consortium on Chicago School Research.
- U.S. Department of Education. (2012). *Application for grants under the TIF general competition* (CFDA No. 84.374A, PR/Award No. S374A120066, Grants.gov Tracking No. GRANT 11189548). Retrieved from <http://www2.ed.gov/programs/teacherincentive/apps2012/s374a120066.pdf>

Appendix A. Descriptive Results

Descriptive Tables for Pretraining Survey

Table A1. Occupational Roles of Administrators, Pretraining Survey Sample

Role	Percentage	Count
Principal	63.9	432
Assistant principal	24.3	164
Instructional director	1.3	9
Central office staff	4.6	31
Teacher	0.3	2
Other	5.6	38

Note. $N = 677$. The counts do not add to 677 because 1 response is missing for this question

Table A2. Number of Years Administrators Served in Same Job Prior to This School Year, Pretraining Survey Sample

Time served	Mean	SD	Min.	Max.
Years	6.0	5.4	0	34

Note. $N = 677$.

Table A3. Grade Levels Administrators Work With Regularly, Pretraining Survey Sample

Grade level	Percentage	Count
Pre-K	37.8	254
K	55.8	375
Grade 1	56.1	377
Grade 2	56.5	380
Grade 3	56.0	376
Grade 4	56.4	379
Grade 5	56.3	378
Grade 6	36.5	245
Grade 7	26.3	177
Grade 8	26.6	179
Grade 9	31.8	214
Grade 10	31.7	213
Grade 11	31.5	212
Grade 12	31.8	214

Note. $N = 677$. Percentages do not add to 100% because some administrators work in multiple grades.

Table A4. Certification Areas of Administrators Who Returned Presurveys

Certification area	Overall (<i>N</i> = 676)	Summer training (<i>N</i> = 529)	School year training (<i>N</i> = 147)
Elementary education	63.5% (429)	65.8% (348)	55.1% (81)
Mathematics	13.8% (93)	14.0% (74)	12.9% (19)
English/language arts	20.71% (140)	21.4% (113)	18.4% (27)
Science	10.36% (70)	11.3% (60)	6.8% (10)
Social studies	17.0% (115)	18.0% (95)	13.6% (20)

Note. Percentages do not add to 100% because some administrators hold multiple certifications. Counts are in parentheses.

Table A5. Certification Areas of Administrators, Pretraining Survey Sample

Certification area	Percentage	Count
I do not have a teaching certification	1.0	7
English language arts	20.8	140
Social studies (any area)	17.1	115
Mathematics	13.8	93
Science (any area)	10.4	70
Vocational/technical education (any area)	2.7	18
Physical education	9.1	61
Special education	12.9	87
Library sciences	0.3	2
Elementary education	63.7	429
Teaching English learners (BCLAD, CLAD, SB1969)	56.2	378
Other (please specify)	12.2	82

Note. *N* = 677. Percentages do not add to 100% because some administrators hold multiple certifications.

Table A6a. Focus Administrators' Weekly Time Allocation by Area

Area of job responsibility	0–2 hours	2–5 hours	5–10 hours	>10 hours
Student discipline	20.5% (8)	30.8% (12)	38.5% (15)	10.3% (4)
Paperwork		7.7% (3)	35.9% (14)	56.4% (22)
Work with stakeholders	2.6% (1)	20.5% (8)	38.5% (15)	38.5% (15)
Individual or small group work with teachers	12.8% (5)	35.9% (14)	30.8% (12)	20.5% (8)
Meeting with other administrators	53.8% (21)	23.1% (9)	17.9% (7)	5.1% (2)
Managing data	23.1% (9)	30.8% (12)	35.9% (14)	10.3% (4)
Activities directly related to classroom instruction	5.1% (2)	41.0% (16)	35.9% (14)	17.9% (7)
Operational duties	2.6% (1)	5.1% (2)	30.8% (12)	61.5% (24)

Note. Frequencies are in parentheses.

Table A6b. Focus Administrators: Important Aspects of Their Job, Other Than Listed

Rank	Important aspect
1	Building capacity
2	Implementing the New Master Plan
3	Instructional program
4	New budget system
5	Operations is not so important to me since I am the instructional director.
6	Single admin site. Everything takes time away from instructional leadership, which I believe is my main job.
7	Student safety from pedophiles in the neighborhood to teenage suicide prevention
8	Supervision of teachers and PD
9	Training for me in Common Core professional development
10	With so much to do, a leader has to prioritize responsibilities. Supervision of instruction will be <i>very</i> important, but the TGDC process is simply too cumbersome and time consuming to be of great importance. If TGDC becomes the focus, then I will be unable to maintain balance between work and home—and ultimately this will make me a much worse leader.

Table A6c. Focus Administrators: Time-Consuming Aspects of Their Job, Other Than Listed

Rank	Time-consuming aspect
1	Disciplinary concerns involving teachers
2	I am not going to do the TGDC with fidelity—and perhaps I will be written up for it. However, my school has 20+ point gains in API for the first 2 years (waiting for new data for my third year) without TGDC, so we will see if compliance or results are what truly matter to district leadership . . .
3	I would like in-service training in Common Core, not just once before the year begins.
4	New budget system
5	Not enough hours in the day
6	WASC, Safe School Plans, Schoolwide Positive Behavior Support Plans and Implementation, EL accountabilities, ISTAR reporting, random metal searching daily, including logs and accountabilities, Master Plan Implementation, IEP mtgs, monthly parent meetings, bimonthly report cards and award assemblies, OEHS, monthly principal meetings off-campus, daily lunch line and supervision . . . I haven't had coffee yet, so I'm just getting started!

Table A7. Administrators' Expectations for Learning in 2012–2013 Prior to Training, Pretraining Survey Sample

Expectations for learning	Not at all	A little	Some	A lot
How observations fit into the TGDC	1.7% (11)	2.4% (16)	30.1% (199)	65.8% (435)
How to score observations more accurately	1.1% (7)	2.3% (15)	23.9% (158)	72.8% (481)
How to manage the TGDC with my other responsibilities	2.3% (15)	2.7% (18)	16.3% (108)	78.7% (521)
How to take good notes when observing	0.8% (5)	5.3% (35)	25.8% (170)	68.2% (450)
How to use MyPGS, the online observation tool	1.2% (8)	4.2% (28)	22.7% (150)	71.8% (474)
How to give teachers feedback based on observed performance	0.6% (4)	3.6% (24)	23.6% (157)	72.1% (479)
How to match teachers to appropriate professional development based on their observed performance	0.8% (5)	3.2% (21)	28.5% (189)	67.5% (447)
How to help people who are resistant to change learn to improve their practice	0.9% (6)	2.7% (18)	18.5% (123)	77.9% (517)
How to effectively manage all the strengths and weaknesses of my staff	1.4% (9)	2.0% (13)	21.1% (139)	75.6% (499)

Note. $N = 677$. Frequencies are in parentheses. TGDC = Teacher Growth and Development Cycle.

Table A8. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data Sources for Improving Instruction, Pretraining Survey Sample

Data source	Not at all useful	Somewhat useful	Mostly useful	Highly useful
Supervisors' observations of teaching	0.2% (1)	4.5% (30)	29.8% (198)	65.6% (436)
Student performance on state standardized tests	1.5% (10)	24.1% (160)	46.3% (308)	28.1% (187)
Parent surveys	5.9% (39)	47.8% (317)	33.6% (223)	12.7% (84)
Teacher self-assessment	0.8% (5)	19.1% (126)	38.7% (255)	41.4% (273)
Professional goal setting and growth plans	0.5% (3)	11.8% (78)	37.5% (248)	50.2% (332)
Student performance on other achievement measures (e.g., benchmark, unit, end of course tests)	0.6% (4)	10.6% (70)	42.6% (282)	46.2% (306)
Student surveys of teachers' performance	4.5% (30)	39.5% (261)	34.6% (229)	21.3% (141)

Note. $N = 677$. Frequencies are in parentheses.

Table A9. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes, Pretraining Survey Sample

Purpose	Not at all useful	Somewhat useful	Mostly useful	Highly useful
Identifying or rewarding strong teachers	3.7% (24)	26.3% (171)	39.4% (256)	30.5% (198)
Teacher improvement/development	0.8% (5)	11.2% (73)	37.4% (244)	50.6% (330)
School improvement	1.1% (7)	14.4% (94)	38.7% (252)	45.8% (298)
Teacher dismissal	5.7% (37)	33.5% (217)	27.2% (176)	33.6% (218)
Creating a common vision of excellent teaching	0.9% (6)	12.9% (84)	33.7% (220)	52.5% (343)
District improvement	2.6% (17)	23.4% (153)	38.4% (251)	35.5% (232)
Provide public accountability	5.2% (34)	29.8% (194)	36.5% (237)	28.5% (185)

Note. $N = 677$. Frequencies are in parentheses.

Table A10. Frequency of Specific Evaluation Practices Prior to 2012–2013, Pretraining Survey Sample

Practice	Never	Occasionally	Most of the time	Always
Conduct informal observations in classrooms on a regular basis	1.5% (10)	14.7% (97)	49.9% (329)	33.8% (223)
Review student work	2.0% (13)	41.6% (273)	39.9% (262)	16.6% (109)
Hold preobservation and postobservation conferences	3.6% (24)	30.7% (202)	40.7% (268)	24.9% (164)
Point out specific strengths in teacher instructional practices in postobservation conferences	2.1% (14)	13.4% (88)	40.1% (264)	44.4% (292)
Point out specific weaknesses in teacher instructional practices in post-observation conference written evaluations	4.0% (26)	19.5% (128)	42.0% (276)	34.6% (227)
Suggest readings or curriculum resources	5.6% (37)	51.1% (337)	31.0% (204)	12.3% (81)
Co-plan before a lesson is delivered	28.5% (187)	53.7% (353)	13.5% (89)	4.3% (28)
Review lesson plans on a regular basis	15.1% (99)	49.6% (326)	26.8% (176)	8.5% (56)

Note. $N = 677$. Frequencies are in parentheses.

Table A11. Focus Administrator Agreement Regarding Observation Practices Prior to Teacher Growth and Development Cycle Training

Observation practice	Agree	Somewhat agree	Somewhat disagree	Disagree
I felt that I was skillful in conducting teacher evaluations.	51.3% (20)	43.6% (17)	5.1% (2)	—
I observe teachers regularly as part of my ongoing work	82.1% (32)	17.9% (7)	—	—
I am able to give teachers useful feedback about their teaching.	59.0% (23)	35.9% (14)	5.1% (2)	—
I feel comfortable observing and evaluating teachers of most grade levels and subjects.	76.9% (30)	20.5% (8)	2.6% (1)	—
I have prior training in teacher observation and evaluation.	59.0% (23)	23.1% (9)	12.8% (5)	5.1% (2)

Note. Frequencies are in parentheses.

Table A12. Frequency of Specific Evaluation Practices Prior to Implementing the Teacher Growth and Development Cycle

Evaluation practice	Never	Occasionally	Most of the time
Conduct informal observations in classrooms on a regular basis			
Overall	1.5% (10)	14.7% (97)	49.9% (329)
Summer	0.2% (1)	12.5% (65)	51.5% (269)
School year	6.6% (9)	23.4% (32)	43.8% (60)
Review student work			
Overall	2.0% (13)	41.6% (273)	39.9% (262)
Summer	1.2% (6)	42.5% (221)	40.4% (210)
School year	5.1% (7)	38.0% (52)	38.0% (52)
Hold preobservation and postobservation conferences			
Overall	3.7% (24)	30.7% (202)	40.7% (268)
Summer	1.5% (8)	31.5% (164)	41.5% (216)
School year	11.7% (16)	27.7% (38)	38.0% (52)
Point out specific strengths in teacher instructional practices in postobservation conferences			
Overall	2.1% (14)	13.4% (88)	40.1% (264)
Summer	0.8% (4)	13.2% (69)	41.5% (216)
School year	7.3% (10)	13.87% (19)	35.0% (48)
Point out specific weaknesses in teacher instructional practices in postobservation written evaluations			
Overall	4.0% (26)	19.5% (128)	42.0% (276)
Summer	2.1% (11)	19.6% (102)	43.1% (224)
School year	11.0% (15)	19.0% (26)	38.0% (52)
Suggest readings or curriculum resources			
Overall	5.6% (37)	51.1% (337)	31.0% (204)
Summer	5.6% (29)	52.7% (275)	30.0% (156)
School year	5.8% (8)	45.3% (62)	35.0% (48)
Co-plan before a lesson is delivered			
Overall	28.5% (187)	53.7% (353)	13.6% (89)
Summer	29.6% (154)	54.2% (282)	13.1% (68)
School year	24.1% (33)	51.8% (71)	15.3% (21)
Review lesson plans on a regular basis			
Overall	15.1% (99)	49.6% (326)	26.8% (176)
Summer	12.3% (64)	51.2% (266)	28.1% (146)
School year	25.6% (35)	43.8% (60)	21.9% (30)

Note. Frequencies are in parentheses.

Table A13a. Administrators' Agreement With Statements About What Makes for Effective Teaching, Pretraining Survey Sample

Statement	Not at all	A little	Some	A lot
The teacher writes detailed lesson plans.	1.2% (8)	14.1% (92)	53.4% (349)	31.2% (204)
The teacher knows what each student is doing.	0.2% (1)	0.8% (5)	13.0% (85)	86.1% (564)
The teacher requires a quiet classroom.	40.8% (266)	41.4% (270)	15.8% (103)	2.0% (13)
Order and discipline come first in the teacher's classroom.	2.3% (15)	15.7% (103)	41.1% (269)	40.8% (267)
Students see the teacher as someone they can look up to.	0.2% (1)	3.1% (20)	25.2% (164)	71.6% (466)
The teacher's students must do more than learn basic facts.	0.0% (0)	0.8% (5)	8.9% (58)	90.4% (591)
The teacher encourages students to express opinions different than his/her own.	0.0% (0)	1.4% (9)	16.9% (111)	81.7% (536)
The teacher engages students in question and answer.	0.2% (1)	4.1% (27)	15.2% (100)	80.5% (528)
The teacher gives students a chance to discuss issues among themselves.	0.2% (1)	0.8% (5)	10.9% (71)	88.2% (576)
The teacher gives students the opportunity to explore subject matter on their own.	0.2% (1)	1.7% (11)	17.6% (115)	80.5% (525)
Other	6.0% (9)	2.7% (4)	5.4% (8)	85.9% (128)

Note. $N = 677$. Frequencies are in parentheses.

Table A13b. Administrators' Agreement With Statements About What Makes for Effective Teaching, Summer Versus School Year Training Participants

Statement	Not at all	A little	Some	A lot
The teacher writes detailed lesson plans.				
Overall	1.2% (8)	14.1% (92)	53.5% (349)	31.2% (204)
Summer	1.4% (7)	13.8% (71)	54.5% (280)	30.4% (156)
School year	0.7% (1)	15.1% (21)	49.6% (69)	34.53% (48)
The teacher knows what each student is doing.				
Overall	0.2% (1)	0.8% (5)	13.0% (85)	86.1% (564)
Summer	0.2% (1)	0.8% (4)	13.2% (68)	85.9% (443)
School year	0.0% (0)	0.7% (1)	12.2% (17)	87.1% (121)
The teacher requires a quiet classroom.				
Overall	40.8% (266)	41.4% (270)	15.8% (103)	2.0% (13)
Summer	41.9% (216)	39.8% (205)	16.7% (86)	1.6% (8)
School year	36.5% (50)	47.5% (65)	12.4% (17)	3.7% (5)
Order and discipline come first in the teacher's classroom.				
Overall	2.3% (15)	15.8% (103)	41.1% (269)	40.8% (267)
Summer	1.9% (10)	14.8% (76)	41.2% (212)	42.1% (217)
School year	3.6% (5)	19.4% (27)	41.0% (57)	36.0% (50)
Students see the teacher as someone they can look up to.				
Overall	0.2% (1)	3.1% (20)	25.2% (164)	71.6% (466)
Summer	0.2% (1)	2.9% (15)	24.1% (124)	72.8% (374)
School year	0.0% (0)	3.7% (5)	29.2% (40)	67.2% (92)
The teacher's students must do more than learn basic facts.				
Overall	0.0% (0)	0.8% (5)	8.9% (58)	90.4% (591)
Summer	0.0% (0)	1.0% (5)	8.8% (45)	90.3% (465)
School year	0.0% (0)	0.0% (0)	9.4% (13)	90.7% (126)
The teacher encourages students to express opinions different than his/her own.				
Overall	0.0% (0)	1.4% (9)	16.9% (111)	81.7% (536)
Summer	0.0% (0)	1.6 (8)	18.2% (94)	80.3% (415)
School year	0.0% (0)	0.7% (1)	12.2% (17)	87.1% (121)
The teacher engages students in question and answer.				
Overall	0.2% (1)	4.1% (27)	15.2% (100)	80.5% (528)
Summer	0.2% (1)	4.3% (22)	15.9% (82)	79.7% (412)
School year	0.0% (0)	3.6% (5)	13.0% (18)	83.5% (116)

Statement	Not at all	A little	Some	A lot
The teacher gives students a chance to discuss issues among themselves.				
Overall	0.2% (1)	0.8% (5)	10.9% (71)	88.2% (576)
Summer	0.2% (1)	1.0% (5)	10.9% (56)	88.0% (453)
School year	0.0% (0)	0.0% (0)	10.9% (15)	89.1% (123)
The teacher gives students the opportunity to explore subject matter on their own.				
Overall	0.2% (1)	1.7% (11)	17.6% (115)	80.5% (525)
Summer	0.0% (0)	2.1% (11)	18.5% (95)	79.3% (407)
School year	0.7% (1)	0.0% (0)	14.4% (20)	84.9% (118)
Other				
Overall	6.0% (9)	2.7% (4)	5.4% (8)	85.9% (128)
Summer	6.6% (8)	3.3% (4)	4.9% (6)	85.3% (104)
School year	3.7% (1)	0.0% (0)	7.4% (2)	88.9% (24)

Note. Frequencies are in parentheses.

Table A14. Administrators' Evaluation of Full-Time K–12 Classroom Teachers in 2011–2012, Pretraining Survey Sample

Evaluation	Mean	SD	Min.	Max.
Highly effective teachers	22.7	20.0	0	100
Effective teachers	35.6	20.0	0	100
Developing teachers	26.5	19.4	0	100
Ineffective teachers	12.4	11.5	0	88

N = 677.

Table A15a. Administrators' Perceptions of Job Efficacy, Pretraining Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I can handle the daily demands of the job.	1.2% (8)	6.9% (45)	37.9% (247)	53.9% (351)
I can focus on the quality of instruction and learning in each classroom.	5.9% (38)	21.8% (141)	42.7% (276)	29.7% (192)
I can shape the operational policies and procedures that are necessary to manage my school.	4.0% (26)	8.6% (55)	43.7% (281)	43.7% (281)
I spend the majority of my time on tasks unrelated to the quality of instruction.	7.7% (50)	19.9% (129)	47.4% (307)	25.0% (162)
I always use the school improvement goals to guide my decision making.	0.9% (6)	7.6% (49)	47.7% (309)	43.8% (284)

Note. *N* = 677. Frequencies are in parentheses.

Table A15b. Administrators’ Perceptions of Job Efficacy, Summer Versus School Year Training Participants

Statement	Disagree	Somewhat disagree	Somewhat agree
I can handle the daily demands of the job.			
Overall	1.2% (8)	6.9% (45)	37.9% (247)
Summer	1.4% (7)	7.4% (38)	39.8% (205)
School year	0.7% (1)	5.2% (7)	30.9% (42)
I can focus on the quality of instruction and learning in each classroom.			
Overall	5.9% (38)	21.8% (141)	42.7% (276)
Summer	5.8% (30)	21.32% (110)	44.2% (228)
School year	6.1% (8)	23.7% (31)	36.6% (48)
I can shape the operational policies and procedures that are necessary to manage my school.			
Overall	4.0% (26)	8.6% (55)	43.7% (281)
Summer	3.5% (18)	7.8% (40)	43.3% (222)
School year	6.2% (8)	11.5% (15)	45.4% (59)
I spend the majority of my time on tasks unrelated to the quality of instruction.			
Overall	7.7% (50)	19.9% (129)	47.4% (307)
Summer	7.4% (38)	19.4% (100)	48.5% (250)
School year	9.0% (12)	21.8% (29)	42.9% (57)

Note. Frequencies are in parentheses.

Table A16. Administrators’ Note-Taking Approach, Pretraining Survey Sample

Approach	Percentage	Count
Summarize instructional activity	15.2	96
Describe instructional activity	46.1	292
Evaluate instructional activity	25.1	159
None of the above	13.6	86

Note. *N* = 677. Descriptive Tables for the Mid-Training Survey

Table A17. Administrators’ Responses to Question 1 on the Mid-Training Survey

Response	Percentage	Count
True	95.6	368
False	4.4	17

Note. *N* = 389.

Table A18. Administrators’ Responses to Question 2 on the Mid-Training Survey

Response	Percentage	Count
True	55.0	209
False	45.0	171

Note. *N* = 389.

Table A19. Administrators’ Responses to Question 3 on the Mid-Training Survey

Response	Percentage	Count
Consistency in the quality of performance observed	5.0	19
Level of knowledge of the profession demonstrated through practice	1.0	4
Amount and level of student cognitive engagement in classroom activities	6.8	26
All of the above	87.2	333

Note. *N* = 389.

Table A20. Administrators’ Responses to Question 4 on the Mid-Training Survey

Response	Percentage	Count
Preobservation conference	93.8	362
Classroom observation	42.5	164
Postobservation	23.8	92

Note. *N* = 389.

Table A21. Administrators’ Responses to Question 5 on the Mid-Training Survey

Response	Percentage	Count
Ineffective	6.6	25
Minimally effective/developing	40.9	156
Effective/proficient	86.4	329
Highly effective/distinguished	55.4	211

Note. *N* = 389.

Descriptive Tables for Posttraining Survey

Table A22. Consistency Between the Training and Administrators' Expectations, Posttraining Survey Sample

Statement	Not at all	A little	Some	A lot
Your personal goals for learning to score observations more accurately	2.5% (16)	7.8% (49)	42.0% (264)	47.6% (299)
Your beliefs about the features of effective instruction	1.4% (9)	6.2% (39)	36.2% (227)	56.1% (352)
Your beliefs about the role of observations in teacher evaluation	1.6% (10)	4.9% (31)	35.5% (223)	58.0% (364)
Your beliefs about the role of administrators in the teacher evaluation process	1.4% (9)	5.4% (34)	34.9% (219)	58.2% (365)

Note. $N = 633$. Frequencies are in parentheses.

Table A23. Administrators' Beliefs About Applying Training to Their Own Schools, Posttraining Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The teachers in my school(s) will be working in a different context than the teachers we observed in training.	32.0% (195)	20.9% (127)	33.3% (203)	13.8% (84)
In some cases, I will be evaluating teachers in grades I am less familiar with.	51.6% (314)	22.9% (139)	16.4% (100)	9.0% (55)
In some cases, I will be evaluating teachers in subjects I am less familiar with.	45.9% (277)	22.2% (134)	21.0% (127)	10.9% (66)
I will not be able to maintain my reliability as a rater over the course of the school year.	49.1% (295)	30.1% (181)	16.8% (101)	4.0% (24)
My existing relationships with teachers in my school will get in the way of my ability to objectively score their observed performance.	72.7% (440)	19.7% (119)	6.6% (40)	1.0% (6)
The observation process we learned will not capture some of what I believe is important.	39.2% (239)	24.8% (151)	25.8% (157)	10.2% (62)
I will not have enough time to do the quality of observation we have learned about this week.	9.2% (56)	14.6% (89)	34.7% (212)	41.6% (254)

Note. $N = 633$. Frequencies are in parentheses.

Table A24. Administrators' Evaluation of Full-Time K–12 Classroom Teachers in 2012–2013, Posttraining Survey Sample

Evaluation	Mean	SD	Min.	Max.
Highly effective teachers	12.5	14.6	0	103
Effective teachers	33.9	21.3	0	100
Developing teachers	38.2	22.8	0	100
Ineffective teachers	10.9	10.1	0	98

Note. $N = 633$.

Table A25. Administrators' Views of Training Outcomes, Posttraining Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I feel confident in being able to identify good evidence properly.	0.5% (3)	4.6% (28)	51.6% (314)	43.3% (263)
I feel confident in my ability to align evidence.	0.8% (5)	7.4% (45)	59.6% (362)	32.1% (195)
I feel confident in my ability to navigate MyPGS, the online observation tool.	1.3% (8)	7.8% (47)	46.7% (283)	44.2% (268)
I feel confident recognizing constructivist learning activities and increased student responsibility when observing practice.	0.5% (3)	7.5% (45)	49.8% (300)	42.3% (255)
I understand the TGDC observation processes and the LAUSD Teaching and Learning rubrics.	0.3% (2)	4.8% (29)	41.8% (253)	53.1% (321)
I understand the concept of 21st-century skills.	1.2% (7)	5.9% (36)	42.1% (255)	50.8% (308)
I understand the concept of cognitive engagement.	0.7% (4)	2.5% (15)	37.9% (229)	58.9% (356)
I understand the distinctions between the levels of performance (highly effective, effective, developing, and ineffective) in the Teaching and Learning rubrics.	0.5% (3)	4.5% (27)	42.1% (252)	52.9% (317)
I feel confident in my ability to match the master raters' scores on training videos.	2.0% (12)	14.2% (85)	64.9% (389)	18.9% (113)

Note. $N = 633$. Frequencies are in parentheses. LAUSD = Los Angeles Unified School District; TGDC = Teacher Growth and Development Cycle.

Table A26a. Administrators’ Views of the Teaching and Learning Framework, Posttraining Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The online system makes the instrument easy to use.	7.6% (46)	17.9% (108)	44.5% (269)	30.0% (181)
The Teaching and Learning Framework covers important domains of teaching.	0.2% (1)	0.5% (3)	23.6% (143)	75.7% (458)
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction.	0.2% (1)	1.7% (10)	30.8% (186)	67.4% (407)
The Teaching and Learning Framework contains the proper amount of detail for observing teaching.	1.8% (11)	7.7% (47)	43.2% (262)	47.3% (287)
Teaching behaviors are adequately specified in this instrument.	1.0% (6)	9.9% (60)	47.3% (287)	41.8% (254)
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.	1.2% (7)	7.3% (44)	46.4% (280)	45.1% (272)
I am comfortable with this observational instrument.	3.5% (21)	16.5% (100)	48.1% (291)	31.9% (193)

Note. *N* = 633. Frequencies are in parentheses.

Table A26b. Administrators' Views of the Teaching and Learning Framework, Summer Versus School Year Training

Participants

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The Teaching and Learning Framework covers important domains of teaching.				
Overall	0.2% (1)	0.5% (3)	23.6% (143)	75.7% (458)
Summer	0.2% (1)	0.6% (3)	24.1% (123)	75.01% (383)
School year	0.0% (0)	0.0% (0)	21.1% (20)	79.0% (75)
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction.				
Overall	0.2% (1)	1.7% (10)	30.8% (186)	67.4% (407)
Summer	0.2% (1)	1.6% (8)	31.0% (158)	67.3% (343)
School year	0.0% (0)	2.1% (2)	29.8% (28)	68.1% (64)
The Teaching and Learning Framework contains the proper amount of detail for observing teaching.				
Overall	1.8% (11)	7.7% (47)	43.2% (262)	47.3% (287)
Summer	2.2% (11)	8.2% (42)	43.1% (220)	46.6% (238)
School year	0.0% (0)	5.2% (5)	43.8% (42)	51.0% (49)
Teaching behaviors are adequately specified in this instrument.				
Overall	1.0% (6)	9.9% (60)	47.3% (287)	41.9% (254)
Summer	0.8% (4)	9.8% (50)	47.2% (241)	42.3% (216)
School year	2.1% (2)	10.4% (10)	47.9% (46)	39.6% (38)
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.				
Overall	1.2% (7)	7.3% (44)	46.4% (280)	45.1% (272)
Summer	1.2% (6)	6.7% (34)	47.2% (240)	45.0% (229)
School year	1.1% (1)	10.6% (10)	42.6% (40)	45.7% (43)
I am comfortable with this observational instrument.				
Overall	3.5% (21)	16.5% (100)	48.1% (291)	31.9% (193)
Summer	4.1% (21)	17.3% (88)	48.9% (249)	29.7% (151)
School year	0.0% (0)	12.5% (12)	43.8% (42)	43.8% (42)
The online system makes the instrument easy to use.				
Overall	7.6% (46)	17.9% (108)	44.5% (269)	30.0% (181)
Summer	9.6% (44)	19.1% (97)	45.0% (229)	27.3% (139)
School year	2.1% (2)	11.6% (11)	42.1% (40)	44.2% (42)

Note. Frequencies are in parentheses.

Table A26c. Administrators' Views of the Teaching and Learning Framework, by Job Role

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The Teaching and Learning Framework covers important domains of teaching.				
Overall	0.2% (1)	0.5% (3)	23.6% (143)	75.7% (458)
Principal	0.4% (1)	0.0% (0)	23.7% (65)	75.9% (208)
Assistant principal	0.0% (0)	1.9% (2)	27.2% (28)	70.9% (73)
Central office staff	0.0% (0)	0.0% (0)	21.4% (3)	78.6% (11)
Instructional director	0.0% (0)	0.0% (0)	20.0% (1)	80.0% (4)
Other	0.0% (0)	0.0% (0)	21.4% (6)	78.6% (22)
NA	0.0% (0)	0.6% (1)	22.1% (40)	77.3% (140)
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction.				
Overall	0.2% (1)	1.7% (10)	30.8% (186)	67.4% (407)
Principal	0.4% (1)	2.2% (6)	30.9% (85)	66.5% (183)
Assistant principal	0.0% (0)	2.0% (2)	40.2% (41)	57.8% (59)
Central office staff	0.0% (0)	7.1% (1)	14.3% (2)	78.6% (11)
Instructional director	0.0% (0)	0.0% (0)	0.0% (0)	100.0% (5)
Other	0.0% (0)	0.0% (0)	35.7% (10)	64.3% (18)
NA	0.0% (0)	0.6% (1)	26.7% (48)	72.8% (131)
The Teaching and Learning Framework contains the proper amount of detail for observing teaching.				
Overall	1.8 (11)	7.7% (47)	43.2% (262)	47.3% (287)
Principal	2.2% (6)	9.1% (25)	44.7% (123)	44.0% (121)
Assistant principal	1.9% (2)	9.7% (10)	48.5% (50)	39.8% (41)
Central office staff	0.0% (0)	7.1% (1)	42.9% (6)	50.0% (7)
Instructional director	0.0% (0)	0.0% (0)	40.0% (2)	60.0% (3)
Other	0.0% (0)	0.0% (0)	32.1% (9)	67.9% (19)
NA	1.6% (3)	6.0% (11)	39.6% (72)	52.7% (96)
Teaching behaviors are adequately specified in this instrument.				
Overall	1.0% (6)	9.9% (60)	47.3% (287)	47.1% (254)
Principal	0.4% (1)	8.3% (23)	47.8% (132)	43.5% (120)
Assistant principal	2.9% (3)	17.5% (18)	46.6% (48)	33.0% (34)
Central office staff	0.0% (0)	7.1% (1)	42.9% (6)	50.0% (7)
Instructional director	0.0% (0)	0.0% (0)	40.0% (2)	60.0% (3)
Other	3.6% (1)	7.1% (2)	35.7% (10)	53.6% (15)
NA	0.6% (1)	8.8% (16)	49.2% (89)	41.4% (75)
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.				
Overall	1.2% (7)	7.3% (44)	46.4% (280)	45.1% (272)

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
Principal	0.7% (2)	7.7% (21)	46.4% (127)	45.3% (124)
Assistant Principal	3.9% (4)	7.8% (8)	52.0% (53)	36.3% (37)
Central Office Staff	0.0% (0)	15.4% (2)	38.5% (5)	46.2% (6)
Instructional Director	0.0% (0)	0.0% (0)	40.0% (2)	60.0% (3)
Other	0.0% (0)	10.7% (3)	28.6% (8)	60.7% (17)
NA	0.6% (1)	5.5% (10)	47.0% (85)	47.0% (85)
The online system makes the instrument easy to use.				
Overall	7.6% (46)	17.9% (108)	44.5% (269)	30.0% (181)
Principal	10.2% (28)	20.8% (57)	43.8% (120)	25.2% (69)
Assistant principal	4.9% (5)	16.5% (17)	43.7% (45)	35.0% (36)
Central office staff	0.0% (0)	21.4% (3)	50.0% (7)	28.6% (4)
Instructional director	0.0% (0)	20.0% (1)	20.0% (1)	60.0% (3)
Other	0.0% (0)	0.0% (0)	53.6% (15)	46.4% (13)
NA	7.2% (13)	16.7% (30)	45.0% (81)	31.1% (56)
I am comfortable with this observational instrument.				
Overall	3.5% (21)	16.5% (100)	48.1% (291)	31.9% (193)
Principal	5.1% (14)	19.7% (54)	45.6% (125)	29.6% (81)
Assistant principal	2.9% (3)	12.6% (13)	50.5% (52)	34.0% (35)
Central office staff	0.0% (0)	7.1% (1)	50.0% (7)	42.9% (6)
Instructional director	0.0% (0)	20.0% (1)	40.0% (2)	40.0% (2)
Other	0.0% (0)	3.6% (1)	50.0% (14)	46.4% (13)
NA	2.2% (4)	16.6% (30)	50.3% (91)	30.9% (56)

Note. Frequencies are in parentheses.

Table A26d. Administrators' Views of the Teaching and Learning Framework, by Instructional Level

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The Teaching and Learning Framework covers important domains of teaching.				
Overall	0.2% (1)	0.5% (3)	23.6% (143)	75.7% (458)
Elementary	0.3% (1)	0.6% (2)	23.1% (75)	75.9% (246)
Middle school	0.5% (1)	1.0% (2)	24.9% (49)	73.6% (145)
High school	0.0% (0)	0.7% (1)	23.5% (32)	75.7% (103)
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction.				
Overall	0.2% (1)	1.7% (10)	30.8% (186)	67.4% (407)
Elementary	0.3% (1)	2.2% (7)	33.0% (107)	64.5% (209)
Middle school	0.5% (1)	3.0% (6)	34.0% (67)	62.4% (123)
High school	0.0% (0)	4.4% (6)	30.1% (41)	65.4% (89)
The Teaching and Learning Framework contains the proper amount of detail for observing teaching.				
Overall	1.8 (11)	7.7% (47)	43.2% (262)	47.3% (287)
Elementary	2.2% (7)	8.0% (26)	44.8% (145)	45.1% (146)
Middle school	1.5% (3)	8.1% (16)	46.0% (91)	44.4% (88)
High school	0.7% (1)	8.0% (11)	44.5% (61)	46.7% (64)
Teaching behaviors are adequately specified in this instrument.				
Overall	1.0% (6)	9.9% (60)	47.3% (287)	41.9% (254)
Elementary	0.6% (2)	9.8% (32)	47.1% (153)	42.5% (138)
Middle school	2.5% (5)	11.6% (23)	46.5% (92)	39.4% (78)
High school	2.9% (4)	11.7% (16)	43.1% (59)	42.3% (58)
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.				
Overall	1.2% (7)	7.3% (44)	46.4% (280)	45.1% (272)
Elementary	1.6% (5)	6.2% (20)	45.5% (146)	46.7% (150)
Middle school	2.0% (4)	12.8% (25)	45.9% (90)	39.3% (77)
High school	1.5% (2)	15.4% (21)	44.1% (60)	39.0% (53)
The online system makes the instrument easy to use.				
Overall	7.6% (46)	17.9% (108)	44.5% (269)	30.0% (181)
Elementary	8.0% (26)	17.0% (55)	45.1% (146)	29.9% (97)
Middle school	7.1% (14)	18.8% (37)	42.6% (84)	31.5% (62)
High school	6.6% (9)	21.3% (29)	42.6% (58)	29.4% (40)
I am comfortable with this observational instrument.				
Overall	3.5% (21)	16.5% (100)	48.1% (291)	31.9% (193)
Elementary	4.6% (15)	15.4% (50)	48.5% (157)	31.5% (102)
Middle school	2.5% (5)	14.1% (28)	50.0% (99)	33.3% (66)
High school	2.2% (3)	14.6% (20)	47.4% (65)	35.8% (49)

Note. Frequencies are in parentheses.

Table A27. Administrators' Views of the Utility of Teacher Growth and Development Cycle Observation Component for Various Purposes, Posttraining Survey Sample

Statement	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers	2.5% (15)	16.8% (101)	41.9% (252)	38.9% (234)
Teacher improvement/ development	0.8% (5)	10.1% (60)	44.1% (263)	45.1% (269)
School improvement	0.8% (5)	13.5% (81)	45.2% (270)	40.5% (242)
Teacher dismissal	7.5% (45)	28.3% (169)	32.4% (194)	31.8% (190)
Increasing the amount of time I can spend on teaching and learning	14.2% (84)	27.2% (161)	36.0% (213)	22.6% (134)
Creating a common vision of excellent teaching	1.8% (11)	8.1% (49)	40.7% (245)	49.3% (297)
District improvement	2.7% (16)	19.3% (116)	44.8% (269)	33.3% (200)

Note. N = 633. Frequencies are in parentheses.

Table A28. Administrators' Views of the Utility of Teacher Growth and Development Cycle Data for Various Purposes, Posttraining Survey Sample

Statement	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers	2.0% (12)	17.6% (105)	44.5% (266)	36.0% (215)
Teacher improvement/development	1.2% (7)	12.5% (74)	46.3% (275)	40.1% (238)
School improvement	1.2% (7)	16.6% (99)	46.7% (279)	35.5% (212)
Teacher dismissal	7.9% (47)	27.2% (162)	32.4% (193)	32.4% (193)
Increasing the amount of time I can spend on teaching and learning	12.7% (75)	29.2% (172)	34.2% (202)	23.9% (141)
Creating a common vision of excellent teaching	1.3% (8)	10.1% (60)	42.3% (252)	46.3% (276)
District improvement	2.9% (17)	19.7% (116)	43.9% (259)	33.6% (198)

Note. N = 633. Frequencies are in parentheses.

Descriptive Tables for End-of-Year Survey

Table A29. Administrators' Views of Training Outcomes, End-of-Year Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I feel confident in my ability to use MyPGS, the online observation tool.	6.3% (25)	17.8% (71)	49.5% (197)	26.4% (105)
I feel confident in being able to collect objective evidence of instruction.	1.5% (6)	6.4% (25)	46.6% (183)	45.5% (179)
I feel confident in my ability to align evidence.	2.5% (10)	11.1% (44)	52.9% (210)	33.5% (133)
I understand the standards, components, and elements in the LAUSD Teaching and Learning rubrics.	1.0% (4)	3.3% (13)	38.2% (151)	57.5% (227)
I understand the differences between the levels of performance (<i>Highly Effective</i> , <i>Effective</i> , <i>Developing</i> , and <i>Ineffective</i>) in the Teaching and Learning rubrics.	1.0% (4)	4.3% (17)	37.8% (150)	56.9% (226)
I feel confident conducting preobservation meetings with teachers.	0.3% (1)	5.8% (23)	36.2% (144)	57.8% (230)
I feel confident conducting postobservation meetings with teachers.	0.5% (2)	5.3% (21)	36.1% (143)	58.1% (230)
LAUSD's training prepared me adequately for my observation work this year.	6.1% (24)	18.8% (74)	53.0% (209)	22.1% (87)

Note. $N = 398$. Frequencies are in parentheses. LAUSD = Los Angeles Unified School District.

Table A30. Number of Formal and Informal Observations Administrators Have Done With Focus Teachers in 2012–2013, End-of-Year Survey Sample

Observation	Mean	SD
Formal observation(s)	2.0	0.9
Informal observation(s)	4.9	7.9

Note. $N = 398$.

Table A31. Total Number of Hours to Complete a Formal Observation Cycle, End-of-Year Survey Sample

Time	Mean	SD
Hours	10.3	10.3

Note. $N = 398$.

Table A32. Administrators’ Impression of the Teaching and Learning Framework, End-of-Year Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The Teaching and Learning Framework covers important domains of teaching.	0.0% (0)	0.0% (0)	19.8% (51)	80.2% (207)
The view of instruction underlying the Teaching and Learning Framework is similar to my own view of instruction.	0.0% (0)	1.2% (3)	27.4% (71)	71.4% (185)
Observing using the Teaching and Learning Framework has helped me think more critically about teaching.	0.4% (1)	5.8% (15)	24.7% (64)	69.1% (179)
The Teaching and Learning Framework is a valuable tool for talking to my teachers about their instruction.	0.0% (0)	3.5% (9)	21.3% (55)	75.2% (194)
The Teaching and Learning Framework will be more likely to improve the quality of teaching in my school than the STULL process did.	3.1% (8)	7.7% (20)	25.1% (65)	64.1% (166)
The Teaching and Learning Framework meets the teaching and learning needs specific to my school.	1.9% (5)	8.9% (23)	36.0% (93)	53.1% (137)
I am comfortable with this observational instrument.	6.3% (16)	14.8% (38)	43.4% (111)	35.5% (91)

Note. $N = 398$. Frequencies are in parentheses.

Table A33. Administrators’ Views of the Utility of Teacher Growth and Development Cycle Observation Component for Various Purposes, End-of-Year Survey Sample

Purpose	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers	3.6% (13)	18.9% (69)	48.9% (179)	28.7% (105)
Teacher improvement/development	1.1% (4)	12.3% (45)	43.3% (159)	43.3% (159)
School improvement	2.7% (10)	17.8% (65)	45.1% (165)	34.4% (126)
Teacher dismissal	10.0% (36)	27.8% (100)	33.1% (119)	29.2% (105)
Increasing the amount of time I can spend on teaching and learning	15.4% (56)	20.9% (76)	43.7% (159)	20.1% (73)
Creating a common vision of excellent teaching	3.6% (13)	10.4% (38)	45.9% (168)	40.2% (147)
District improvement	5.2% (19)	22.0% (80)	43.3% (157)	29.5% (107)
Preparing teachers for working with the Common Core State Standards	4.1% (15)	23.0% (84)	43.7% (160)	29.2% (107)

Note. $N = 398$. Frequencies are in parentheses.

Table A34. Administrators' Perceptions of Job Efficacy, End-of-Year Survey Sample

Perception	Disagree	Somewhat disagree	Somewhat agree	Agree
I can handle the daily demands of the job.	3.6% (13)	14.0% (51)	41.8% (152)	40.7% (148)
I can focus on the quality of instruction and learning in each classroom.	7.1% (26)	26.6% (97)	41.5% (151)	24.7% (90)
I can shape the operational policies and procedures that are necessary to manage my school.	5.5% (20)	19.1% (69)	45.2% (163)	30.2% (109)
I always use the school improvement goals to guide my decision making.	2.5% (9)	6.7% (24)	41.7% (150)	49.2% (177)

Note. $N = 398$. Frequencies are in parentheses.

Table A35. Administrators' Implementation of Teacher Growth and Development Cycle (TGDC) Training, End-of-Year Survey Sample

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I will be able to implement the TGDC with all eligible teachers.	23.4% (85)	23.1% (84)	37.1% (135)	16.5% (60)
I will be able to handle the daily demands of my job.	9.6% (35)	21.8% (79)	39.4% (143)	29.2% (106)
I will be able to focus on the quality of instruction and learning in each classroom.	7.2% (26)	21.1% (76)	47.9% (173)	23.8% (86)
I will be able to carry out the operational policies and procedures that are necessary.	5.5% (20)	15.2% (55)	48.1% (174)	31.2% (113)

Note. $N = 398$. Frequencies are in parentheses.

Table A36. Challenges Associated With Observing and Evaluating Specific Teacher Groups

Teacher group	A lot more challenging	More challenging	Slightly more challenging	Not more challenging
Very experienced teachers	17.9% (7)	28.2% (11)	28.2% (11)	25.6% (10)
Teachers older than the principal	12.8% (5)	20.5% (8)	28.2% (11)	38.5% (15)
Teachers with whom a principal has had a conflict	30.8% (12)	30.8% (12)	23.1% (9)	15.4% (6)
Strongest teachers in the school	2.6% (1)	23.1% (9)	25.6% (10)	48.7% (19)
Teachers who are personally not liked	7.7% (3)	20.5% (8)	15.4% (6)	56.4% (22)
Teachers outside a principal's subject matter expertise	5.1% (2)	20.5% (8)	38.5% (15)	35.9% (14)
Special education teachers	15.4% (6)	0.0% (0)	38.5% (15)	46.2% (18)
Teachers working at grade levels less familiar to the principal	5.1% (2)	15.4% (6)	30.8% (12)	48.7% (19)
Teachers who have a different style than the observing principal	0.0% (0)	17.9% (7)	43.6% (17)	38.5% (15)

Note. Frequencies are in parentheses.

Table A37. Focus Administrators: Comments on Teacher Growth and Development Cycle (TGDC)—Open-Ended Responses

Rank	Comment
1	Add to the platform a component to sort out and manage informal observations. Informal observations are the equivalent of progress monitoring in tracking student performance. In essence, the predictability value of formal observations, as a measure of teacher effectiveness, decreases in the absence of informal observations. Formal observations would come to confirm what a collection of informal observations would have established long before any formal observation. A process that includes informal observations decreases the high-stakes nature of a few formal observations. A model of growth and development should be more compatible with frequent informal observations by virtue of its frequency. It has the potential to add to the effectiveness of feedback that can make a difference almost instantaneously. Thank you for your consideration.
2	Although I am timidly looking forward to participating in the TGDC process, I am greatly concerned about the time it takes to complete just one person and the fact that the process is a bit cumbersome to complete. It was a struggle this year and I was unable to complete the process with my teacher, due to taking over this school late in the year. This year I have up to this point a total of 11 teachers to evaluate. My only other concern is whether or not the district's technology system can handle everyone accessing the system at the same time and continue to operate without freezing up and/or shutting down. For example, when I certify payroll for my staff, I strive to get it done no later than the day before. Because when everyone is trying to certify payroll at the same the time the system constantly kicks me off several times during the process, and I tend to lose data and am unable to retrieve it. I trust that these issues have been addressed before hundreds of personnel are asked to do the impossible with nothing. :-)
3	Be patient with teachers. Make sure they have many opportunities to learn.
4	Because I am the only administrator in a small school setting, I cannot share my TGDC tasks and responsibilities with anyone else.
5	Collaboration with the teachers' union is crucial in TGDC's full implementation. I believe next year will be a real test of the full implementation as it will be replacing actual teacher evaluation.
6	Don't change the platform mid-year; it was difficult to figure out the tech nuances in addition to task completion.
7	TGDC has been an excellent professional growth experience, and I am grateful for the new skills I have gathered throughout the process.
8	I do believe that the greatest drawback to the TGDC is the amount of time needed to complete the process. Single site administrators would be hard-pressed to perform all the operational tasks required and then complete full TGDC cycles.
9	I had problems with the software, logging into the server.
10	I requested TGDC assistance with finding a teacher to work with from another school. It was promised from two individuals at the ESC level, but it was not delivered even though I asked a few times. My questions regarding [being] matched with a teacher that was promised were ignored. I must state that I became very disenchanted with the TGDC process, and I did not complete the three cycles this year because of this.
11	It is so complicated and not user friendly. UCLA uses a similar system of observation and notation, but I do not know what type of platform they use to enter their data. I love the idea, I enjoyed the process, but felt like a complete failure with the actual software. I spent a tremendous amount of time with my TGDC teacher, and he and I both learned so much about his teaching practice and my observation abilities and skill set, but he had as much trouble with the program as I did, and he is tech-savvy. In addition, neither of us is a naysayer, and we entered this endeavor with open minds.
12	Little support. Only mandates
13	Streamline the process and increase administrator norms, and it has immense potential to positively impact student learning. The current process with the current administrator norms ensures schools will not be able to implement with fidelity and results will be mixed at best. Good luck!!!
14	The discussions with the teacher I was observing were an integral part of shortening my learning curve. The frequent instructional rounds using the T&LF with my instructional director and support team were beneficial in using and becoming more familiar with the T&LF. The debrief after each grade level instructional round was great for my learning and reflection. Also, my participation in going through the National Board Certified Teacher

process baptized me in the rigor and reflection required for aligning classroom evidence to the T&LF. That the TGDC teacher and I had gone through this process previously was a plus. However, carving out the time to complete the TGDC process was a real challenge.

15 There is too much reliance on “scripting” teacher and student behavior during the observation. Administrators lacking the typing speed to significantly script are at a disadvantage and may not adequately provide observation feedback, though they may be strong instructional leaders.

16 Think it will continue to improve as we become more familiar with it

17 To be able to properly implement the TGDC will require lots of time. The system has undergone some changes to facilitate the online process; however, it will still demand an extraordinary amount of time to complete a full observation cycle. Administrators will not be able to effectively use the process if they have a large amount of teachers to evaluate. Consideration also needs to be taken for the following: a below standard evaluation. How does a below standard evaluation change the process? Thank you for conducting this study!

Note. ESC = Education Service Center; T&LF = Teaching and Learning Framework; UCLA = University of California, Los Angeles.

Table A38a. Certification Results (by Occupational Role)

Certification results	Overall		Objectivity		Alignment		Representation		Accuracy	
	Principal	Assist.	Principal	Assist.	Principal	Assist.	Principal	Assist.	Principal	Assist.
Certified with distinction	0%	0%	0%	1%	0%	0%	0%	0%	6%	4%
Certified	10%	5%	78%	73%	45%	40%	32%	27%	31%	25%
Preliminarily certified	80%	84%	18%	21%	43%	48%	56%	57%	35%	37%
Not yet certified	10%	11%	4%	6%	12%	12%	12%	16%	27%	35%

Note. Assist. = assistant principal. Principal, $N = 700$; assistant principal, $N = 320$.

Table A38b. Certification Results (by Grade Level)

Certification results	Overall		Objectivity		Alignment		Representation		Accuracy	
	Elem.	Sec.	Elem.	Sec.	Elem.	Sec.	Elem.	Sec.	Elem.	Sec.
Certified with distinction	0%	0%	0%	0%	0%	0%	0%	0%	7%	5%
Certified	11%	8%	80%	72%	47%	41%	34%	28%	32%	32%
Preliminarily certified	80%	82%	16%	23%	40%	48%	56%	56%	35%	36%
Not yet certified	9%	10%	3%	5%	13%	11%	10%	16%	26%	28%

Note. Elem. = elementary ($N = 471$); Sec. = secondary ($N = 264$).

Appendix B. Analyses

Table B1. Pretraining Survey Measures Used to Predict Training Satisfaction

Survey measure	Presurvey items	Mean response (1–4)		
		Overall	Summer training	School year training
Importance of student growth data in teacher evaluation	6f	3.34 (0.69)	3.31 (0.70)	3.46 (0.64)
Importance of teacher evaluation data	7a–7g	3.12 (0.66)	3.08 (0.67)	3.27 (0.58)
Perception of job manageability	11a–11d	3.15 (0.46)	3.15 (0.46)	3.16 (0.45)
High expectations for observation training	5a–5i	3.67 (0.44)	3.65 (0.45)	3.74 (0.36)
Belief in student-centered instruction	9f–9j	3.82 (0.29)	3.81 (0.30)	3.86 (0.23)

Table B2. Regression Coefficients Predicting Training Satisfaction Using Pretraining Measures of Beliefs and Expectations

Item	Model 1: Assessment of learning	Model 2: Confidence in conducting observations	Model 3: Training met expectations
Pretraining beliefs/expectations			
Importance of student growth data in teacher evaluation	.05(.05)	–.01 (.04)	.02 (.06)
Importance of teacher evaluation data	.23*** (.06)	.14** (.06)	.17*** (.05)
High expectations for observation training	.10* (.05)	–.02 (.06)	.22** (.08)
Perception of job manageability	.05 (.05)	.11*** (.02)	–.03 (.04)
Belief in student-centered instruction	–.14* (.08)	.04 (.06)	.06 (.05)
Administrator characteristics			
Elementary	–.11 (.09)	–.03 (.05)	.03 (.05)
Middle	–.09 (.07)	–.09* (.05)	–.04 (.04)
High	.03 (.10)	.09 (.05)	.08 (.05)
Principal	–.04 (.09)	–.07 (.10)	–.11 (.10)
Assistant principal	.06 (.11)	–.07 (.09)	–.17 (.13)
Training type: Summer (vs. school)	–.06 (.06)	–.04 (.07)	–.03 (.08)

Note. $N = 408$. Standard errors are in parentheses. To account for shared experiences within training sites, we included robust standard errors.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Table B3. Predicting Certification Using Pre- and Posttraining Measures of Beliefs and Expectations

Item	Model 1: Overall	Model 2: Objectivity	Model 3: Alignment	Model 4: Representation	Model 5: Accuracy
Presurvey measures					
Importance of teacher evaluation data	-0.18*(0.11)	0.20 (0.24)	-0.10 (0.08)	-0.11 (0.26)	-0.04 (0.12)
High expectations for observation training	0.37 (0.35)	-0.23 (0.35)	0.16 (0.31)	0.13 (0.21)	-.08 (0.26)
Perception of job manageability	-0.15 (0.27)	0.15 (0.22)	-0.17 (0.23)	-0.23** (0.09)	-0.34 (0.40)
Postsurvey measures					
Assessment of learning	-0.25 (0.21)	-0.72*** (0.27)	-0.14 (0.11)	-0.06 (0.20)	-0.30* (0.18)
Training met expectations	0.20(0.24)	0.27 (0.20)	0.39 (0.24)	0.39** (0.18)	-0.10 (0.25)
Confidence in ability to conduct observations	0.27 (0.48)	0.27 (0.24)	-0.16 (0.34)	0.26 (0.28)	-0.02 (0.15)
Administrator characteristics					
Elementary	0.49 (0.33)	0.30 (0.54)	0.040 (0.21)	0.07 (0.37)	-0.03 (0.35)
Middle	-0.15 (0.38)	-0.63 (0.49)	0.61** (0.29)	0.11 (0.31)	0.06 (0.31)
High	0.43 (0.43)	0.61 (0.59)	-0.53 (0.43)	-0.14 (0.32)	-0.35 (0.54)
Principal	0.02 (0.41)	0.35 (0.57)	0.21 (0.31)	-0.28 (0.45)	0.82* (0.47)
Assistant principal	-0.41 (0.80)	0.53 (0.73)	-0.23 (0.49)	0.03 (0.18)	0.47 (0.66)
Training type: Summer (vs. school)	0.89 (0.23)	0.42 (0.93)	0.13 (0.27)	1.48 (0.42)	-0.04 (0.29)

Note. $N = 349$. All certification variables are treated as dichotomous outcomes (3 = CD or C; 2 = PC; 1 = NYC). All coefficients presented as odds ratios, representing the likelihood of becoming certified in a given domain. An odds ratio greater than 1.00 indicates an increased likelihood, whereas an odds ratio less than 1.00 indicates a decreased likelihood.

* $p < .10$. ** $p < .05$. *** $p < .01$.

Table B4. Pearson Correlation Among Certification Measures

Measure	Mean	SD	Min.	Max.	Objectivity	Alignment	Represent	Accuracy
Objectivity	2.74	0.54	1	4	1			
Alignment	2.39	0.67	1	4	0.26***	1		
Represent	2.27	0.65	1	4	0.34***	0.40***	1	
Accuracy	2.15	0.85	1	4	0.071**	0.12***	0.14***	1

* $p < .10$. ** $p < .05$. *** $p < .01$.

Table B5. Element-Level Reliability Statistics (Full Sample)

Element	Category	% Exact match (true score)	Mean deviation	% Exact match (other administrators)
Standard 1: Planning and Preparation		42	0.24	42
1d1	Analysis ad use of assess. data for planning	61	-0.32	44
1d3	Standards-based learning activities	52	0.01	39
1d4	Purposeful instructional groups	65	-0.25	47
1e1	Lesson and unit structure	30	0.53	44
1e2	Aligns with instructional outcomes	41	0.36	39
1e3	Criteria and standards	30	0.60	38
1e4	Design of formative assessments	17	0.76	44
Standard 2: Designing Coherent Instruction		69	-0.11	52
2a1	Teacher interactions with students	69	-0.02	53
2a3	Classroom climate	66	-0.02	49
2b2	Expectations for learning and achievement	69	-0.34	51
2c1	Management of routines, procedures, and transitions	70	0.03	55
2d2	Monitoring and responding to student behavior	72	-0.19	55
Standard 3: Standards-Based Learning Activities		48	0.12	44
3a1	Communicating the purpose of the lesson	20	0.70	48
3b1	Quality and purpose of questions	38	0.49	39
3b2	Discussion techniques	65	-0.18	47
3c1	Standards-based projects, activities, and assignments	63	-0.36	46
3c2	Purposeful and productive grouping of students	60	-0.20	42
3d1	Assessment criteria	10	1.08	38
3d3	Feedback to students	67	-0.21	49
3e1	Responds and adjusts to meet student needs	59	-0.36	41
Standard 5: Professional Growth		43	0.23	32
5a2	Use of reflection to inform future instruction	43	0.23	32

Table B6. Element-Level Reliability Statistics (by Job Role)

Element	Statistic	Asst. principal % exact match	Principal % exact match
Standard 1: Planning and Preparation		41	45
1d1	Analysis and use of assess. data for planning	67	60
1d3	Standards-based learning activities	53	56
1d4	Purposeful instructional groups	68	66
1e1	Lesson and unit structure	23	35
1e2	Aligns with instructional outcomes	36	45
1e3	Criteria and standards	28	32
1e4	Design of formative assessments	13	20
Standard 2: Designing Coherent Instruction		67	72
2a1	Teacher interactions with students	66	73
2a3	Classroom climate	62	70
2b2	Expectations for learning and achievement	73	70
2c1	Management of routines, procedures, and transitions	64	75
2d2	Monitoring and responding to student behavior	68	75
Standard 3: Standards-Based Learning Activities		47	50
3a1	Communicating the purpose of the lesson	15	22
3b1	Quality and purpose of questions	32	43
3b2	Discussion techniques	69	67
3c1	Standards-based projects, activities, and assignments	63	67
3c2	Purposeful and productive grouping of students	65	61
3d1	Assessment criteria	8	11
3d3	Feedback to students	62	70
3e1	Responds and adjusts to meet student needs	60	60
Standard 5: Professional Growth		37	47
5a2	Use of reflection to inform future instruction	37	47

Table B7. Element-Level Reliability Statistics (by Instructional Level)

Element		Statistic	Elem. % exact match	Sec. % exact match
Standard 1: Planning and Preparation			45	41
1d1	Analysis and use of assess. data for planning		64	59
1d3	Standards-based learning activities		55	53
1d4	Purposeful instructional groups		68	64
1e1	Lesson and unit structure		34	25
1e2	Aligns with instructional outcomes		45	40
1e3	Criteria and standards		32	30
1e4	Design of formative assessments		18	16
Standard 2: Designing Coherent Instruction			71	70
2a1	Teacher interactions with students		71	70
2a3	Classroom climate		70	64
2b2	Expectations for learning and achievement		71	73
2c1	Management of routines, procedures, and transitions		72	72
2d2	Monitoring and responding to student behavior		74	72
Standard 3: Standards-Based Learning Activities			50	48
3a1	Communicating the purpose of the lesson		22	15
3b1	Quality and purpose of questions		41	38
3b2	Discussion techniques		68	67
3c1	Standards-based projects, activities, and assignments		65	66
3c2	Purposeful and productive grouping of students		60	66
3d1	Assessment criteria		11	8
3d3	Feedback to students		70	66
3e1	Responds and adjusts to meet student needs		61	60
Standard 5: Professional Growth			44	46
5a2	Use of reflection to inform future instruction		44	46

Note. Elem. = elementary; Sec. = secondary.

Appendix C. Year 1 Study Instruments

Los Angeles Unified School District Administrator Pretraining Survey

During this week of training, you will be learning a great deal about the observation component of the Educator Growth and Development Cycle (EGDC). We would like to learn about your expectations going into the training. This will help us learn how training can be improved and how people with different experiences understand the training. Your individual responses to these items will be confidential and your identity will not be revealed to anyone. The survey should take no more than 10 minutes.

If you have any questions or have difficulty completing the online survey, please contact Dr. Courtney Bell at cbell@ets.org (609.273.6328) or Brian Lucas at brian.lucas@lausd.net (213.241.3444).

Background Information

1. Which position best describes your primary role in the district?

- ☐ Principal
- ☐ Assistant Principal
- ☐ Instructional Director
- ☐ Central Office Staff
- ☐ Teacher
- ☐ Other (Please specify)

2. PRIOR to this school year (2013–2014), how many years did you serve in the role you referenced in Question #1 in THIS OR ANY OTHER DISTRICT?

3. In which of these grades do you primarily work? *Select all that apply*

- ☐ Pre-K ☐ Grade 6

- | | |
|----------------------------------|-----------------------------------|
| <input type="checkbox"/> K | <input type="checkbox"/> Grade 7 |
| <input type="checkbox"/> Grade 1 | <input type="checkbox"/> Grade 8 |
| <input type="checkbox"/> Grade 2 | <input type="checkbox"/> Grade 9 |
| <input type="checkbox"/> Grade 3 | <input type="checkbox"/> Grade 10 |
| <input type="checkbox"/> Grade 4 | <input type="checkbox"/> Grade 11 |
| <input type="checkbox"/> Grade 5 | <input type="checkbox"/> Grade 12 |

4. In what subject area(s) are you certified to teach? *Select all that apply*

- | | |
|--|--|
| <input type="checkbox"/> I do not have a teacher certification | <input type="checkbox"/> Physical Education |
| <input type="checkbox"/> English Language Arts | <input type="checkbox"/> Special Education |
| <input type="checkbox"/> Social Studies (any area) | <input type="checkbox"/> Library Sciences |
| <input type="checkbox"/> Mathematics | <input type="checkbox"/> Elementary Education |
| <input type="checkbox"/> Science (any area) | <input type="checkbox"/> Teaching English Learners (BCLAD, CLAD, SB1969) |
| <input type="checkbox"/> Vocational/Technical Education (any area) | |
| <input type="checkbox"/> Other (please specify) | |

Expectations for Training and Prior Experiences

Questions 5–13 ask about your expectations for training and prior experience conducting teacher evaluation.

5. Over the course of the 2013–2014 school year, to what extent would you like to learn the following? *Select one answer for each row*

Statement	Not at all	A little	Some	A lot
How observations fit into the TGDC				
How to score observations more accurately				
How to manage the TGDC with my other responsibilities				
How to take good notes when observing				
How to use MyPGS, the online observation tool				
How to give teachers feedback based on observed performance				
How to match teachers to appropriate professional development based on their observed performance				
How to help people who are resistant to change learn to improve their practice				
How to effectively manage all the strengths and weaknesses of my staff				

6. The TGDC draws on each of the data sources below. To what degree do you think that each of the following TGDC data sources will be useful for improving instruction?

Select one answer for each row

Data source	Not at all useful	Somewhat useful	Mostly useful	Highly useful
Supervisors' observations of teaching				
Student performance on state standardized tests				
Parent surveys				
Teacher self-assessment				
Professional goal setting and growth plans				
Student performance on other achievement measures (e.g., benchmark, unit, end of course tests, etc.)				
Student surveys of teachers' performance				

7. To what degree do you think that the TGDC data will be useful for the following purposes?

Select one answer for each row

Purpose	Not at all useful	Somewhat useful	Mostly useful	Highly useful
Identifying or rewarding strong teachers				
Teacher improvement/development				
School improvement				
Teacher dismissal				
Creating a common vision of excellent teaching				
District improvement				
Provide public accountability				

8. When evaluating classroom instruction prior to implementing the TGDC, how often did you do each of the following? *Select one answer for each row*

Evaluation	Never	Occasionally	Most of the time	Always
Conduct informal observations in classrooms on a regular basis				
Review student work				
Hold preobservation and postobservation conferences				
Point out specific strengths in teacher instructional practices in postobservation conferences				
Point out specific weaknesses in teacher instructional practices in postobservation conferences written evaluations				
Suggest readings or curriculum resources				
Co-plan before a lesson is delivered				
Review lesson plans on a regular basis				

9. To what degree does each of the following reflect your views of what makes for effective teaching? *Select one answer for each row*

Statement	Not at all	A little	Some	A lot
The teacher writes detailed lesson plans				
The teacher knows what each student is doing				
The teacher requires a quiet classroom				
Order and discipline come first in the teacher's classroom				
Students see the teacher as someone they can look up to				
The teacher's students must do more than learn basic facts				
The teacher encourages students to express opinions different than his/her own				
The teacher engages students in question and answer				
The teacher gives students a chance to discuss issues among themselves				
The teacher gives students the opportunity to explore subject matter on their own				
Other				

Other (please specify)

10. For the 2012–2013 school year, what percentage of FULL-Time K-12 CLASSROOM TEACHERS you worked with would you put in the following categories, based on your overall opinion of their TEACHING PERFORMANCE? (Please note: Percentages should sum to 100. If no teachers fit the category, please write “0”)

Highly Effective teachers

Effective teachers

Developing teachers

Ineffective teachers

11. In your current position, to what extent do you agree with each of the following statements?

Select one answer for each row

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I can handle the daily demands of the job				
I can focus on the quality of instruction and learning in each classroom				
I can shape the operational policies and procedures that are necessary to manage my school				
I spend the majority of my time on tasks unrelated to the quality of instruction				
I always use the school improvement goals to guide my decision making				

12. Below are some examples of notes other administrators have taken when observing the same video of classroom instruction where a third-grade teacher is teaching the story Chrysanthemum by Kevin Henkes. When observing classroom instruction for the purpose of evaluating teachers, which of the following best reflects your note-taking approach?



Summarize instructional activity.

Administrator notes: “The teacher is teaching the book Chrysanthemum by Kevin Henkes. She begins with questions about what students know about the book and after they respond, she moves on to reading the book. Many different students respond to teachers’ questions.”



Describe instructional activity

Administrator notes: “The teacher asks: What is Chrysanthemum about? A blond female student in the third row holds her hand up for approx. 45 seconds and then answers the teacher’s question: ‘Chrysanthemum is a story about self-esteem.’ The teacher responded:


‘Chloe has an interesting idea about the main idea of the story. What do other children think?’ There was then a pause of approx. 20 seconds before 5–6 children raised their hands.”

☐ Evaluate instructional activity

Administrator notes: “The teacher is working on comprehension using Chrysanthemum. She begins with basic questions and positively reinforces students’ responses. Students are watching the teacher and appear interested in the book. This seems to be an appropriate book choice for these students.”

☐ None of the above

13. If you would like to make any additional comments in response to any of the questions in this survey, please include them below:



Thank you for your time!

Los Angeles Unified School District Administrator Day 3 Survey

Thank you for taking a few minutes to complete this survey about today's professional development session. We appreciate your feedback.

1. True or False: The teacher should come to the pre-conference prepared to present an overview of the lesson and information about his/her planning.

- ☐ True
- ☐ False

2. True or False: Only the administrator prepares performance "ratings" to share at the postobservation conference.

- ☐ True
- ☐ False

3. Choose the Best Answer: The primary difference between Level II and Level III performance on the LAUSD T & L Rubrics is:

- ☐ Consistency in the quality of performance observed
- ☐ Level of knowledge of the profession demonstrated through practice
- ☐ Amount and level of student cognitive engagement in classroom activities
- ☐ All of the above

4. Check All That Apply: Evidence of planning and preparation is a primary focus of data collection in the following events in the observation process.

- ☐ Pre-Observation Conference
- ☐ Classroom Observation
- ☐ Post-Observation

5. Check All That Apply: At which level(s) of performance within the rubrics must there be evidence that teachers have included learning experiences designed to facilitate students' construction of knowledge in their classroom instruction?

- ☐ Ineffective
- ☐ Minimally Effective/Developing
- ☐ Effective/Proficient
- ☐ Highly Effective/Distinguished

Part II: Feedback on the Activities From Day 3

6. To what degree do you agree with each of the following statements?

Statement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
As a result of the training, I have a better understanding of the TGDC observation process					
As a result of the training, I have a better understanding of the distinctions between the levels of performance in the Teaching and Learning Framework rubrics					
The views of teaching and learning in the Teaching and Learning Framework match my own.					

7. Please rate the following characteristics of today's workshop:

Characteristic	Poor	Fair	Average	Good	Excellent
Content					
Pacing					
Quality					

Part III: Forward Planning Based on Today's Activities and Discussions

8. What (if any) additional support would you find helpful in building your understanding of the TGDC observation process?

9. How did the priorities influence your evidence collection and your scoring? How is this similar to or different from your current practice?

10. What implications for your current work emerged from today's workshop?

11. What questions or concerns (lingering, anticipatory, burning, etc.) would you like addressed during the remainder of these professional development sessions?

12. Additional comments:

Los Angeles Unified School District Administrator Posttraining Survey

In this survey, we ask you to reflect on your training experience over the past week. This will help us learn how training can be improved and how people with different experiences understand the training. As a reminder, your individual responses to these items will be confidential and your identity will not be revealed to anyone. The survey should take no more than 10 minutes.

If you have any questions or have difficulty completing the online survey, please contact Dr. Courtney Bell at cbell@ets.org (609.273.6328) or Brian Lucas at brian.lucas@lausd.net (213.241.3444).

1. When reflecting on your training, to what extent do you think you learned the following?

Select one answer for each row

Item	Not at all	A little	Some	A lot
How observations fit into the TGDC				
How to score observations more accurately				
How to manage the TGDC with my other responsibilities				
How to take good notes when observing				
How to use MyPGS, the online teacher evaluation tool				
How to give teachers feedback based on observed performance				
How to match teachers to appropriate professional development based on their observed performance				
How to help people who are resistant to change learn to improve their practice				
How to effectively manage all the strengths and weaknesses of my staff				

2. To what extent was the training consistent with your expectations going into training?

Select one answer for each row

Expectation	Not at all	A little	Some	A lot
Your personal goals for learning to score observations more accurately				
Your beliefs about the features of effective instruction				
Your beliefs about the role of observations in teacher evaluation				
Your beliefs about the role of administrators in the teacher evaluation process				

3. When thinking about applying your training to the evaluation of teachers in your school, to what degree do each of the below represent areas of concern?

Select one answer for each row

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The teachers in my school(s) will be working in a different context than the teachers we observed in training				
In some cases, I will be evaluating teachers in grades I am less familiar with				
In some cases, I will be evaluating teachers in subjects I am less familiar with				
I will not be able to maintain my reliability as a rater over the course of the school year				
My existing relationships with teachers in my school will get in the way of my ability to objectively score their observed performance				
The observation process we learned will not capture some of what I believe is important				
I will not have enough time to do the quality of observation we have learned about this week				

4. For the 2012–2013 school year, what percentage of FULL-TIME K-12 CLASSROOM TEACHERS you worked with would you put in the following categories, based on your overall opinion of their TEACHING PERFORMANCE? (Please note: Percentages should sum to 100. If no teachers fit the category, please write “0”)

Highly Effective Teachers

Effective Teachers

Developing Teachers

Ineffective Teachers

5. As a result of participating in this training, to what degree do you agree with the following training outcomes? *Select one answer for each row*

Statements	Disagree	Somewhat disagree	Somewhat agree	Agree
I feel confident in being able to identify good evidence properly				
I feel confident in my ability to align evidence				
I feel confident in my ability to navigate MyPGS, the online observation tool				
I feel confident recognizing constructivist learning activities and increased student responsibility when observing practice				
I understand the TGDC observation processes and the LAUSD Teaching and Learning rubrics				
I understand the concept of 21st century skills				
I understand the concept of cognitive engagement				
I understand the distinctions between the levels of performance (<i>Highly Effective</i> , <i>Effective</i> , <i>Developing</i> , and <i>Ineffective</i>) in the Teaching and Learning rubrics				
I feel confident in my ability to match the master raters' scores on training videos				

6. What is your impression of the Teaching and Learning Framework? *Select one answer for each row*

Impression	Disagree	Somewhat disagree	Somewhat agree	Agree
The online system makes the instrument easy to use				
The Teaching and Learning Framework covers important domains of teaching				
The view of instruction underlying the Teaching and Learning Framework is similar to my view of instruction				
The Teaching and Learning Framework contains the proper amount of detail for observing teaching				
Teaching behaviors are adequately specified in this instrument				
The Teaching and Learning Framework meets the teaching and learning needs specific to my school				
I am comfortable with this observational instrument				

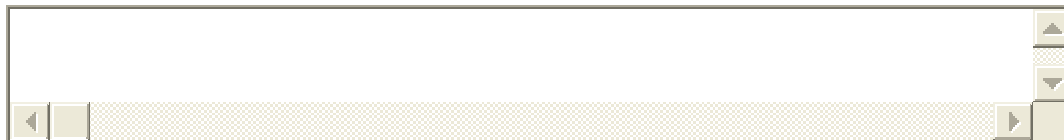
7. To what degree do you think the observation component of TGDC will be useful for the following purposes? *Select one answer for each row*

Purpose	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers				
Teacher improvement/development				
School improvement				
Teacher dismissal				
Increasing the amount of time I can spend on teaching and learning				
Creating a common vision of excellent teaching				
District improvement				

8. To what degree do you think that TGDC data (including all components) will be useful for the following purposes? *Select one answer for each row*

Purpose	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers				
Teacher improvement/development				
School improvement				
Teacher dismissal				
Increasing the amount of time I can spend on teaching and learning				
Creating a common vision of excellent teaching				
District improvement				

9. If you would like to make any additional comments in response to any of the questions in this survey, please include them below:



Thank you for your time!

Los Angeles Unified School District Focus Administrator Training Questionnaire

Welcome to the training questionnaire. In this questionnaire, we are interested in learning about your ideas and beliefs about teaching evaluation as well as your experiences in learning to use the observation parts of the TGDC. Your responses to these questions will be confidential and your identity will not be revealed to anyone. Please be as honest as possible. Your candid responses will help us better understand and improve the system for all administrators. Thanks for your participation!

1. Please indicate the degree to which you agree with each of the statements below.

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I am skillful in conducting teacher evaluations				
I observe teachers regularly as part of my ongoing work				
I am able to give teachers useful feedback about their teaching				
I feel comfortable observing and evaluating teachers of most grade levels and subjects				
I have prior training in teacher observation and evaluation				
The current training has prepared me adequately to do teacher observation for evaluation purposes				

2. When evaluating classroom instruction prior to TGDC, to what extent did you do each of the following?

Statement	Not at all	A little	Some	A lot
Rely on my own sense of high-quality teaching				
Refer to district policy				
Use an observation protocol				
Use the observed teacher's goals as a framework for observation				
Consider the needs of the students and teachers specific to the local context				
Conduct frequent but brief classroom walk-throughs				
Use educational research to inform my evaluations				
Keep a detailed script when observing a lesson				

3. To what degree do you agree with each of the statements about the TGDC observation system?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
My PGS, the online teacher evaluation system, makes the Teaching and Learning Framework easy to use				
The Teaching and Learning (T&L) Framework, covers important areas of teaching				
The T&L Framework's view of instruction is like mine				
The T&L Framework contains the proper amount of detail for observing teaching				
Teaching behaviors are adequately specified in the T&L Framework				
The T&L Framework meets the teaching and learning needs specific to my school				
I am very familiar with the T & L Framework.				
I have been able to provide input in the development of how teachers are evaluated in LAUSD.				

4. In your experience, to what degree are the following groups of teachers more challenging than most teachers to observe and evaluate?

Statement	Not more challenging	A little more challenging	Somewhat more challenging	A lot more challenging
Very experienced teachers				
Teachers older than the principal				
Teachers with whom a principal has had a conflict				
Strongest teachers in the school				
Teachers who are personally not liked				
Teachers outside a principal's subject matter expertise				
Special education teachers				
Teachers working at grade levels less familiar to the principal				
Teachers who have a different style than the observing principal				
Other (please specify)				

5. In your opinion, how important are the following principal actions for supporting teachers to do their best teaching?

Statement	Not at all important	A little important	Moderately important	Very important
Be a supportive and caring colleague				
Hold teachers accountable to external standards				
Hold rigorous standards for student learning				
Observe and provide feedback				
Release teachers to do peer observations				
Encourage collaborative professional learning communities				
Promote collective schoolwide measures of success				
Make appropriate professional development workshops available				
Encourage and support teachers' individual teaching styles				
Foster a shared vision for excellent education				
Maintain a well-run and orderly school environment				
Shield teachers from difficult parents				
Review and comment on lesson plans				
Other (please specify)				

6. For the 2011–2012 school year, what percentage of your time at work was spent doing teacher evaluation activities? (Please give a number between 0 and 100% of your time.)

7. For the 2011–2012 school year, what percentage of your time at work was spent conducting teacher observations? (Please give a number between 0 and 100% of your time.)

8. To what extent do you agree with the following statements?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I am satisfied with my work				
I am happy with my current school assignment				
I am glad to be working in my district				
I think about working outside of public education				
If I could do it all over again, I'd choose the same profession				
I'd consider moving to a different position in my district				
I'd like to stay in education but move to a different school or school district				
I am able to accomplish the things I hope to at work				
I feel I have control over what I do at work				
I am satisfied with my life				

9. How much time do you spend on these areas of your job PER WEEK?

Job area	0–2 hours	2–5 hours	5–10 hours	More than 10 hours
Student discipline				
Paperwork				
Work with stakeholders				
Individual or small group work with teachers				
Meeting with other administrators				
Managing data				
Activities directly related to teaching and learning				
Operational duties				

10. If you would like to make any additional comments in response to any of the questions in this survey, please include them below:

Thank you for your time!

Understanding Consequential Assessment Systems for Teachers Focus Administrator Training Interview Protocol

Thank you for making the time to do this interview with me. We are hoping to use this as a time to learn about how you think about your role in the evaluation process, how you view good teaching, and what your concerns are going into this new TGDC process. If it is alright with you, I'd like to record our interview so that I can make sure I've gotten everything you've said. I cannot always type as fast as people can talk. I will delete the recording as soon as I've typed up my notes from our conversation.

1. I understand you are a [job role]. Can you tell me a little bit about the most important parts of being a [job role]?
2. As a [job role], you already have experience having evaluation conversations with teachers or others that report to you. How do you approach those conversations?
3. As a [job role] do you have to evaluate anyone? Who is that? How do you approach those conversations?
4. PROBE: What are your goals? What do you do when you have to discuss things you anticipate will be difficult? What do you do when the conversation is with someone who is difficult in one way or another?
5. Training is being done with videos from other places with teachers you do not know. As you move to observing in your own building or area, what challenges do you expect to face in conducting the observations that are a part of TGDC?
6. PROBE: If there come times when you are unsure of the content, what do you think you will do? If you are unsure of the developmental level of the students, what might you do? Languages other than the ones you speak?
7. Everyone has different definitions of good teaching. From your perspective, what do good teachers do?
8. Some people have told us that there are "intangibles" of good teachers. What are your intangibles (or things that go beyond the basic job description)?
9. PROBE: (Ask about each of the ones named) Why does [intangible] matter to you?
10. What strengths or resources will help you complete the observation part of TGDC?

11. These observations are just one piece of the larger TGDC system. What kinds of changes, positive or negative, do you expect to come out of TGDC implementation?
12. PROBES: What, if any, positive outcomes do you expect to arise? What are your main concerns about TGDC implementation?

Los Angeles Unified School District Mid-Year Focus Administrator Questionnaire 1

Thank you for agreeing to complete this brief questionnaire about your experience to date with the TGDC observation process. The purpose of this questionnaire is to see where you are in the process, how much time the TGDC activities have taken so far, and the areas with which you would like additional support.

The questionnaire should take about 5 minutes to complete. As with the previous questionnaires in this project, your responses will be confidential and your identity will not be revealed to anyone. Your candid responses will help us better understand and improve the training and ongoing supports available to administrators.

1. How did you select a teacher to observe for the TGDC pilot this year?

I haven't yet selected a teacher.

I asked him/her

I asked for volunteers and he/she volunteered

He/she approached me about wanting to participate

Other: _____

2. Please fill in the number of informal and formal TGDC observations you have done with your teacher or teachers in 2012–2013.

Formal Observation(s):

Informal Observation(s):

3. Including all aspects of the process (e.g., communicating with the teacher, your preparation, the pre- and post-conferences), roughly how long does it take to do a formal observation?
(Please use numbers to indicate the amount of time; e.g., 6 hours, 45 minutes)

4. If you have completed at least one formal TGDC observation, please indicate the degree to which you agree with the following statements:

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The TGDC training adequately prepared me to conduct pre- and post-conferences with my focal teacher.				
The TGDC training adequately prepared me to conduct observations with my focal teacher.				
The TGDC data will be useful for teacher improvement/development.				
I am concerned that I will not have enough time to do the quality of observation that we learned about in our training.				

5. We are trying to plan a visit to L.A. this spring to conduct a second round of interviews. We would like to come around the time that you are planning on doing an observation (formal or informal). What months might be better for you? (Please check all months that could work for you.)

February, March, April, May, June

Thank you!

Los Angeles Unified School District Bimonthly Focus Administrator Questionnaire 2

Hi again!

Thank you for your continued commitment to giving LAUSD feedback on the implementation of TGDC. This questionnaire should take no longer than 5 minutes to complete. The main focus of this questionnaire is on your experiences this year conducting informal observations (the short observations you have done with your focus teacher that do NOT include the pre- and postobservation steps).

As with the previous questionnaires, your responses to these questions will be confidential and your identity will not be revealed to anyone. Your candid responses will help us better understand and improve the training process for administrators.

1. Please list how many informal observations you have done so far this year with your focus teacher. (If you have not completed any informal observations yet, please write "0.")

Informal Observations(s):

2. Including all aspects of the process (e.g., communicating with the teacher, visiting the classroom, documenting the observation), roughly how long does it take to do an informal observation? (Please use numbers to indicate the amount of time; e.g., 0 hour, 45 minutes.)

Hours:

Minutes:

3. To what degree do you agree with the following statements about informal observations?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
Informal observations provide information that is helpful to me in managing my staff.				
Informal observations give me enough information to support the teacher's development.				
Informal observations are helpful but cannot replace formal observations.				
I use informal observations differently than I use formal observations.				
I need both informal and formal observations to support teacher development.				

4. Related to conducting the TGDC, please rank the top three things you most need training on to prepare for the 2013–2014 school year. (Please assign only one #1, one #2, and one #3 ranking.)

Statement	1	2	3
Scoring teachers' lesson plans and other preobservation materials			
Giving teachers feedback on their formal observations			
Understanding the differences among levels on the rubric (<i>Highly Effective, Effective, Developing, and Ineffective</i>)			
Understanding the core concepts in the Teaching and Learning Framework			
Understanding the whole TGDC observation process			
Aligning evidence			
Strategies for working with teachers who I believe can improve but may be somewhat resistant			
Matching master raters' scores on training videos			
Managing the transition from the pilot to full implementation of the TGDC observations			

5. Comments: If you would like to offer any other comments on TGDC or our research on it, please put them in the text box below.

THANK YOU for your time!!

Los Angeles Unified School District Bimonthly Focus Administrator Questionnaire 3

Hi again!

This is your last monthly survey this year. It should take no longer than 10 minutes to complete. The main focus of this questionnaire is on your experiences this year in using the TGDC and how it fits into all of the work you do with the teachers with whom you work. As with the previous questionnaires, your responses to these questions will be confidential and your identity will not be revealed to anyone. Your candid responses will help us better understand how people use and learn to use the TGDC system.

1. Looking forward to next year, how *important* do you think each of the following aspects of your job will be?

Aspect	Not at all important	A little bit important	Somewhat important	Very important
The TGDC Observation Process				
Implementation of the Common Core State Standards				
Operational Tasks (e.g., interacting with parents, physical plant, safety, etc.)				
Other: (please specify)_____				

2. Looking forward to next year, how *time consuming* do you think each of the following aspects of your job will be?

Aspect	Not at all time consuming	A little bit time consuming	Somewhat time consuming	Very time consuming
The TGDC Observation Process				
Implementation of the Common Core State Standards				
Operational Tasks (e.g., interacting with parents, physical plant, safety, etc.)				
Other: (please specify)_____				

3. Taking account of all the teachers/principals you supervise, please indicate what proportion falls in each category.

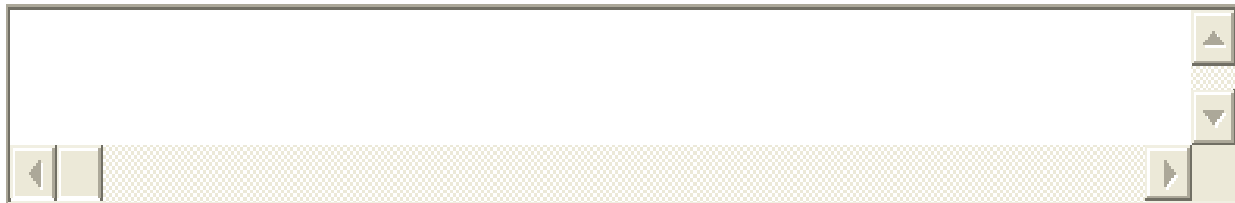
Self-starter, constant learner, needs very few suggestions for improvement	
Learner, responds positively to suggestions for improvement	
Able to learn, does not take suggestions for improvement easily	
Does not take suggestions for improvement at all, should not be teaching	
Does not meet any of these descriptions	

Note: The total should add up to 100. Please do not include the percent symbol (%) in your answer.

4. To what extent have you drawn on each of the following individuals or resources to learn more about the TGDC this year?

Statement	Not at all	A little	Some	A lot
My instructional director (ID)				
My teaching and learning coordinator (TLC)				
Other principals/assistant principals				
My TGDC training materials				
Central administration staff				
Other (please specify)				

5. Comments: If you would like to offer any other comments on the TGDC or our research on it, please put them in the text box below



THANK YOU for your time!!

Los Angeles Unified School District Focus Administrator Think-Aloud Interview Protocol

INTERVIEWER: Thank you for agreeing to help us understand how administrators learn to score. Today, we are going to try to understand how you think about domains 2 and 3 of the Teaching and Learning Framework. In order to do that, we are going to watch a video together with a transcript in front of us, and then listen to you as you assign scores to what you saw.

Some of what we do today will be similar to what you do in your usual observations and some of it will be different. I will describe the whole process to you so you have a big picture view and then I'll answer any questions you have. First, we will watch the 10 minute video together. You can follow along on the transcript if you would like. You can take additional notes, but you do not have to. After the video is over, I will ask you to align the notes and then give a score to just 3 of the focus elements. While you do all of that, I will be quiet, watch, and listen. I won't interrupt you but I will write down places where I don't understand something you have said. I will wait and ask you my questions at the very end.

I mentioned that we call this a think-aloud protocol. Think-aloud protocols try to understand how someone does something by listening to them talk out loud as they do the thing. So in the parts where I ask you to speak out loud, you just need to say the thoughts that are going through your mind. When we have done these think-alouds with other principals some have said they like it, others have said it feels strange. Whatever your reaction, please do not worry about what you are saying. There are no silly things to say. We know so little about how administrators do the work of observing so it is really helpful to just hear people think.

Would it help if I give you some examples of the kinds of things you might say? [pause for response] When you are aligning evidence you might say something like "I am trying to find the spot where there is evidence of the teacher's questioning." or "I'm trying to decide if I want to use this quotation as evidence for questioning and discussion or for using assessment." When you are assigning scores you might say something like "I don't think the teacher did a very good job asking questions. They were mostly low level."

After you are done scoring, I will ask you some questions about what you said. And then we'll be all done. I think this should take us about an hour. Does all that make sense? [pause for response] Do you have any questions before we get started? [pause for response]

Before we begin, you should know that this is a 7th grade “regular” mathematics classroom and the video starts part way through the homework review. Then you see about five minutes of the first activity he does with the students.

BEGIN VIDEO. VIDEO IS 10 MINUTES

AFTER VIDEO IS OVER:

INTERVIEWER: Ok, the next step is for you to align and score as you normally would. You will use the electronic copy of the transcript to cut and paste at least one piece of evidence for each of the three elements. If you would like to include more than one piece of evidence you can, but you don’t have to. When you decide on your overall score, you should use all the evidence, but you do not need to include all the evidence on the scoring sheet. Again, I will be quiet, watch and take notes as you talk out loud. OK? Questions?

Before you start, sometimes it is hard for me to get everything down; would it be ok if I record you so I can get complete notes? I will delete the recording once I have completed my notes.

So that we can learn a little bit more about how principals are thinking, we are asking everyone some clarifying questions about the elements you scored. So I will start with those:

1. For *Expectations for Learning & Achievement* you gave the teaching an xx. Can you tell me a little bit more about why you gave it xx and not a xx-1 or an xx+1?
 - a. Probe: How did you decide what level of expectations the teacher had?
 - b. Probe: How did you decide how many students the teacher had those expectations for?
2. For *Quality/Purpose of Questions* you gave the teaching an xx. Can you tell me a little bit more about why you gave it xx and not a xx-1 or an xx+1?
 - a. Probe: How did you decide what kinds of questions were being asked?
 - b. Probe: How could you tell if the teacher was differentiating the questions for students?
3. For *Standards-Based Projects, Activities & Assignments* you gave the teaching an xx. Can you tell me a little bit more about why you gave it xx and not a xx-1 or an xx+1?

- a. Probe: How did you decide whether the activities were rigorous and appropriate for the students?
 - b. Probe: How did you decide whether students were cognitively engaged by the activities?
4. [Ask other clarifying questions]
5. How does what we just did compare with what you have been doing with the teacher at your school this year?
6. Which was the hardest? Why?
7. Which element was the easiest to score? Why?
8. What's your overall assessment of this excerpt of teaching?
 - a. (If needed) What makes you say so?
 - b. At what point did you arrive at that impression in your mind?
9. Are there any other things I should know about how you think about scoring before we end today?

Focus Administrator Spring Interview Protocol

Subordinate/supervisor issues

1. When you conduct observations with your teachers what do you hope your teachers gain from the evaluations? What do you want to gain from evaluations?
2. There are at least two goals to the TGDC observations – evaluating teachers and helping teachers improve. How do you juggle those 2 goals when you conducting observations?
 - a. PROBE: What strategies do you use to make the teacher more comfortable in your pre- and post-conversations with them?
3. In what ways has the TGDC process changed your interactions with your focus teacher?
 - a. PROBE: This year you only worked with one teacher, next year you will work with more. In what ways might the TGDC process change your interactions with those teachers?

4. We have been told that principals have different relationships with the union in their buildings. What role does the union in your building play in how you do evaluations?
5. We know that there was not really any conversation last summer during training about how to have evaluation conversations. But you have those kinds of conversations all the time. Can you remember how you learned to have evaluation conversations with teachers/principals?
6. What do teachers need in order for the TGDC to be useful?

Training

1. What parts of the TGDC have been the easiest for you to learn? Why? Which have been the hardest?
 - a. PROBE: Are there any insights, shortcuts, tips you have for using the TGDC that you might share with others?
2. Now that you have used the TGDC a bit more, what do you think are its strengths and weaknesses?
3. What do you think would most help principals learn to use the TGDC well?
4. We know that the instructional directors and TLCs across the city have been providing different kinds of learning opportunities. What types of supports have been offered to you? Have you taken advantage of them? How?
5. Taking a step away from the TGDC, if you were in charge of allocating money to support the learning of principals and IDs in the district, what would you prioritize as the most important?

Perception of the Evaluation Policy

1. Can you help me understand what the new evaluation policy will be? What sources of information will be collected in the evaluation of teachers and how will this information be combined?

2. What do you think of the policy? PROBE: Strengths of the policy? Challenges associated with implementation?

Los Angeles Unified School District Administrator End-of-Year Survey

In this survey, we ask you to reflect on your experience with the TGDC observation process this school year. This will help us learn how training can be improved and how we can better support personnel throughout the school year. Your individual responses to these items will be confidential and your identity will not be revealed to anyone. The survey should take no more than 10–15 minutes.

1. To what degree do you agree with the following statements?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I feel confident in my ability to use MyPGS, the online observation tool				
I feel confident in being able to collect objective evidence of instruction				
I feel confident in my ability to align evidence				
I understand the standards, components, and elements in the LAUSD Teaching and Learning rubrics				
I understand the differences between the levels of performance (<i>Highly Effective</i> , <i>Effective</i> , <i>Developing</i> , and <i>Ineffective</i>) in the Teaching and Learning rubrics				
I feel confident conducting preobservation meetings with teachers				
I feel confident conducting postobservation meetings with teachers				
LAUSD's training prepared me adequately for my observation work this year.				

2. We understand some people were not able to work with a focus teacher for a range of reasons. Please check ALL the boxes that apply to your experiences in 2012–2013.

- ☐ I was not able to work with a focus teacher this year
- ☐ I conducted formal observations with my focus teacher
- ☐ I conducted informal observations with my focus teacher
- ☐ I did not conduct any observations with my focus teacher
- ☐ It was difficult to find a focus teacher to work with this year

3. Please fill in the number of formal and informal observations you have done with your focus teacher in 2012–2013.

Formal observation(s)

Informal observation(s)

4. Including all aspects of the process (e.g., your preparation, the classroom observation itself, recording and tagging evidence, entering evidence into MyPGS, the pre- and post-conferences), roughly how long does it take to complete one formal observation cycle?

Note: Please round up to the nearest hour

5. Given your answer to #4, was the time you spent too little, too much, or an appropriate amount of time for a formal observation?

- ☐ Too little time
- ☐ An appropriate amount of time
- ☐ Too much time

6. Now that you have completed your pilot year using the TGDC, what is your impression of the Teaching and Learning Framework?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
The Teaching and Learning Framework covers important domains of teaching				
The view of instruction underlying the Teaching and Learning Framework is similar to my own view of instruction				
Observing using the Teaching and Learning Framework has helped me think more critically about teaching				
Teaching and Learning Framework is a valuable tool for talking to my teachers about their instruction				
The Teaching and Learning Framework will be more likely to improve the quality of teaching in my school than the STULL process did				
The Teaching and Learning Framework meets the teaching and learning needs specific to my school				
I am comfortable with this observational instrument				

7. To what degree do you think the observation component of TGDC will be useful for the following purposes?

Statement	Not at all useful	A little useful	Mostly useful	Very useful
Identifying or rewarding strong teachers				
Teacher improvement/development				
School improvement				
Teacher dismissal				
Increasing the amount of time I can spend on teaching and learning				
Creating a common vision of excellent teaching				
District improvement				
Preparing teachers for working with the Common Core State Standards (CCSS)				

8. In your current position, to what extent do you agree with each of the following statements?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I can handle the daily demands of the job				
I can focus on the quality of instruction and learning in each classroom				
I can shape the operational policies and procedures that are necessary to manage my school				
I always use the school improvement goals to guide my decision making				

9. Thinking about the 2013–2014 school year, to what extent do you agree with each of the following statements?

Statement	Disagree	Somewhat disagree	Somewhat agree	Agree
I will be able to implement the TGDC with all eligible teachers				
I will be able to handle the daily demands of the job				
I will be able to focus on the quality of instruction and learning in each classroom				
I will be able to carry out the operational policies and procedures that are necessary to manage my school				

10. If you would like to make any additional comments in response to any of the questions in this survey, please include them below:

Thank you for your time!