Fall 2022

System-level Leaders' Local Control of Teacher Supervision and Evaluation Under Every Student Succeeds Act

Douglas Wieczorek

Israel Aguilar

Ian M. Mette

System-level Leaders' Local Control of Teacher Supervision and Evaluation Under Every Student Succeeds Act

Douglas Wieczorek, PhD
Associate Professor
Iowa State University
School of Education
Ames, IA

Israel Aguilar, PhD
Assistant Professor
University of Texas, Rio Grande Valley
Department of Organization and School Leadership

Ian Mette, PhD
Associate Professor
Educational Leadership
College of Education and Human Development
University of Maine
Orono, ME

Abstract

This article investigates school system leaders' influence and control over local teacher supervision and evaluation systems (TSES) guided by the United States' (U.S.) Every Student Succeeds Act (ESSA). Using qualitative, textual document analysis methods, we analyzed 50 states' ESSA policies to determine the extent to which local education agencies have flexible TSES authority granted by state and federal guidelines. The study findings indicate that a majority of U.S. state-level policies mandate standardized TSES tools and processes at the local district level. In order to optimally meet students' and teachers' needs, we recommend that systems-level leaders prioritize community driven visions for teachers' professional growth and student learning while maintaining appropriate responsiveness to state and federal educational policy requirements.

Key Words

district leadership, teacher supervision, teacher evaluation, educational policy, instructional leadership, local control

The United States' (U.S.) *Every Student Succeeds Act* (ESSA, 2015, Title II, Part A, sec. 2002., 129 STAT. 1920) codifies federal guidance and support for states to develop respective state-level principal development and teacher supervision and evaluation systems (TSES). The latest ESSA policy provides individual states with the flexible authority to determine unique state and local-district level TSES processes to improve teachers' instruction, professional development, and student learning outcomes.

Despite this shift towards increased federal policy flexibility, little scholarly or practitioner research has analyzed the extent that states' ESSA policy language and legal codes provides local school districts with the influence, agency, and control over TSES in their school communities and respective buildings (Edgerton, 2019; Gagnon, Hall, & Marion, 2017; Kim & Sun, 2021). Furthermore, researchers have not analyzed how states developed TSES policies in response to federal ESSA guidelines to determine the potential impacts on local school district systems leaders' practices such as superintendents and directors.

The purpose of this study is to analyze how all 50 U.S. states developed their TSES in response to ESSA guidelines to determine district leaders' influence and control over TSES processes. This study answers two research questions: 1) During ESSA implementation, what state-level TSES policy requirements and procedures govern local school districts' TSES development and implementation; and 2) What are the implications of these state-level TSES requirements and procedures for systems-level leadership practices at the local district level?

This paper is significant to systems-level leadership scholarship and practice for two reasons; first, our study findings provide systems-level leaders with research-based evidence and examples of how state- and federal-level ESSA policies have continued to leverage bureaucratic control over local districts' TSES autonomy. Second, we discuss and propose leadership actions that systems-level leaders can take to diffuse the effects of these policies on leadership practices in local contexts and recapture local community agency and control in the TSES process.

Analytical Framework

In this article, we argue that the role of district leaders' instructional supervision and evaluation leadership, especially during periodical federal and state policy transitions, is an understudied area that impacts teachers', teacher leaders', and principals' effectiveness to supervise and evaluate instruction aimed to increase student learning. Our analysis indicates that in the majority of U.S. states under ESSA, states continued to leverage significant levels of administrative and procedural control over district-level TSES processes, with evidence of continued state-level policy conflation of supervision and evaluative structures and processes.

The study's findings provide evidence of how ESSA has rhetorically provided a measure of state and local flexibility, but ultimately, the federal and state-level policy responses continue to prevent a significant level of local district authority and control over teacher development, growth, and evaluation of effectiveness. Based on our evidence, local school systems leaders may need to develop strategies to advocate for locally relevant visions and processes for professional growth

and instructional performance, respectively that reflects local community priorities and strategic goals.

Teacher supervision and evaluation practices

As a central focus of this study, research on teacher supervision and evaluation has established these separate, yet closely related leadership processes as essential to students' learning and experiences by supporting teacher development, growth, and professional performance, respectively (Mcintyre & McIntyre, 2020; Tuytens, Devos, & Vanblaere, 2020).

However, there are complex, overlapping, and sometimes conflicting interactions between these two conceptual areas of research and practice, which have been documented by scholars and leaders in practice (McGhee, 2020). Researchers have long advocated for clearly defined supervision roles and procedures that provide teachers with supportive coaching and opportunities to improve practice in safe professional spaces, utilizing peers, colleagues, and principals to provide constructive feedback (Author Three, 2017).

However, there is also an organizational need to evaluate and assess teachers' performance on regular cycles, for the purposes of contract renewal, termination, and as part of a regular performance assessment or clinical observation schedule, which can support but also be contrary to the purpose and spirit of supervision processes (Hazi, 2018). In this study, we have drawn from each conceptual area of supervision and evaluation practices to determine how states have outlined the processes and priorities for teacher growth and performance across contexts. From this analysis we determine potential implications for systems level

leaders' practices to engage with teachers and leaders in TSES in effective and collaborative ways.

ESSA era teacher supervision and evaluation policy

The hierarchical nature of the U.S. educational policy and funding system exhibits intersecting spheres of influence and control at the local community district, state, and federal levels (Debray-Pelot & McGuinn, 2009; Kirst, 1984; Koppich & Esch, 2012; Schneider & Saultz, 2020).

To varying degrees, TSES policy control in the U.S. has alternated between and overlapped among policymakers within the federal government, the individual states, and leaders in local districts, particularly as a result of federal policy iterations implemented through *No Child Left Behind* (NCLB, 2002), *Race to the Top* (RTTT, 2014), and *Every Student Succeeds Act* (ESSA, US PL 114-95, 2015).

Recent research on TSES under ESSA has investigated the changing policy directions between the state and federal level, finding evidence of how states de-emphasized value-added models shifted to a greater focus on teacher development and growth (Close, Amrein-Beardsley, & Collins, 2020; Pauffler, King, & Zhu, 2020).

Research findings also demonstrate that TSES do not manifest as mutually exclusive components of district- and building-level school leadership practices that effectively support both teachers' professional development and document evaluation of teachers' effective practices (Ford & Hewitt, 2020; Lane, 2019). Although these recent policy level findings are important, local community district responses and effects have implications for systems-level leadership

regarding how leaders plan, develop, and implement TSES for teacher growth and development, and ultimately, increasing student learning and achievement.

Systems-level TSES leadership and impacts on stakeholders

During these successive U.S. federal policies from NCLB to ESSA, there is evidence of how systems-level educational leadership perceptions and responsibilities have increased pressures on teacher and leaders in response to changing standards, guidance, and codes at the local, state and federal level regarding teacher quality, effectiveness, supervision, and evaluation (Koppich & Esch, 2012; Pauffler, King, & Zhu, 2020).

This study of TSES policy models is framed by the hierarchical structure, control, and history of education policy pressures exerted at the federal, state, and local levels, specifically building on recent scholar-practitioner research that has documented how teachers and leaders continue to interpret, negotiate, and manage expectations for teacher effectiveness within and across states (Close, Amrein-Beardsley, & Collins, 2018, 2020; Edgerton, 2019; Ford & Hewitt, 2020; Kim & Sun, 2021; Lane, 2020; Pauffler, King, & Zhu, 2020). With respect to district leaders' TSES practices, this continues to be an understudied area of scholarship (Donaldson, Mavrogardato, Youngs, Dougherty, Al Ganem, 2021; Stosich, 2020), which provides opportunities for further examination.

Drawing on evidence from research and practice, we assert that local, systemic control of TSES processes is a more effective, efficient, and contextually relevant approach to support and monitor teachers' professional development and instructional performance. We advocate that scholars, practitioners, and policymakers continue to negotiate, develop,

and implement TSES that draw upon local systems leaders' expertise, resources, and directives to improve teacher performance and student outcomes.

Methods

We completed a qualitative document analysis (Bowen, 2009; Silverman, 2000) of 50 states' TSES policy documents and related ESSA compliance documents to investigate the level of control afforded to local districts over teachers' instructional supervision and evaluation practices. We purposefully selected and analyzed states' policy language, codes, structures, and processes to determine if states' ESSA policies potentially limited local systems leaders' capacity to develop and implement district driven TSES priorities and practices.

We produced a national-level comparative analysis, determining how states across the U.S. have structured TSES policies at the state- and local levels. This study is an adapted extension of an interim 30-state analysis previously completed by the authors and distinguishes from the previous study and related extant research by utilizing a full corpus of national data and applying a multi-level, policy-driven analytical perspective focused on the potential impacts of state policy on local school systems leaders' practices.

Data sources and collection

From February 2019 through May 2020, we collected and archived over 300 publicly available documents from state departments of education (DOE) and state legislature websites, which included legislative documents, state legal codes, DOE regulations, memos on TSES implementation, archived state-level presentations, state legislative technical reports, and white papers which analyzed, described, or outlined TSES policies.

In some cases, we also emailed or called state departments of education to clarify or to obtain information from websites that were not readily accessible due to inactive web links or lack of digital archiving and access. We retained these documents and records of our analysis in a shared digital cloud storage system.

Data analysis and interpretation

We restricted our investigation to analyze state-level policies, related legal documents, and technical reports specifically related to two administrative functions nested within our analytical framework of state- and local leadership and control of instructional improvement: 1) teacher supervision as a means of teachers' formative professional development, and 2) teacher evaluation as a summative assessment of teachers' performance. We utilized a four-step methodological data analysis process described as follows:

(1) Subjectivity and trustworthiness.

First, we articulated our collective subjectivity statement as scholars who care deeply about the intersection of leadership, policy, practice, and research. Thus, applying our beliefs and scholarly knowledge as means to improve school-based practices served as a working heuristic and a starting point of comparison, debate, and reflection. Moreover, we each believe that teachers and administrators are the primary drivers of school improvement and innovation at the local district and school levels. We are critical of reform efforts that enforce top-down mandates and specify the use of particular models, frameworks, and ratings systems as a way to determine local teacher effectiveness.

To mitigate our biases, we engaged in collaborative dialogue and peer-check protocols to ensure that we individually and collectively applied consistent, evidence-based interpretations of policy language and document content. For the duration of the study, we met monthly to consult one another, peer check our methods processes, and complete our analysis. We met regularly to discuss and account for terminology and concepts embedded in policy rationales, and practices not aligned with discussions reflected across the scholarly literature. We considered how states and local districts articulated their respective policies and implementation plans, and we held ourselves accountable to a standard of achieving interpretive and analytical consensus during the analysis process.

(2) Data coding.

Next, we identified specific terminology and language about each state's TSES policies and feedback models, which we organized according to our two primary categories, supervision and evaluation, and these guiding concepts became the initial categories of qualitative codes. For example, for teacher supervision we identified and coded policy and practice terms, which were descriptive of legal code requirements, guidance, processes, purposes, or the materials of practice related to formative professional support in order to improve teachers' instruction.

We identified and included these terms in this category: *instructional supervision, teacher reflection, self-reflection, self-evaluation, coaching, professional development, portfolio development, peer-to-peer conferencing, teacher growth, teacher leadership, mentoring, and teacher improvement.* For teacher evaluation, we identified and coded policy and practice terms related to professional summative evaluation of teachers' instruction and judgment of their professional performance. We identified and included these terms in this category:

instructional evaluation, summative evaluation, teacher ratings, evaluation labels or categories, and applications of rubrics or ratings scales for the purposes of providing measured judgments of teachers' performance.

We identified, but initially withheld, the final categorization of terms and items that we identified as instructional frameworks and tools related to the practice of teachers' instruction. Examples of these items included various instructional frameworks provided by individual state education departments, state teacher unions, or third-party vendors such as Danielson, Marzano, or similar producers of instructional materials intended to guide teachers' instructional practices, assessment, and professional behaviors. This last group of items required an additional level of analytical coding and categorical scrutiny central to our framework, which we describe next.

(3) Data interpretation.

Third, aligned with our framework, at this stage we analyzed how state and local districts constructed TSES systems, what tools were appropriated for TSES purposes, and how the authority for TSES systems were allocated across these systems, respectively. To do this, we compared our respective analytical perceptions of the data, working to reconcile terms and practices that potentially intersected both supervision and evaluation categories.

We employed a cross-case check of each other's state- and local-level language analysis to ensure consistency and trustworthiness regarding how we individually interpreted the intersecting data. We realized TSES models were not only being used at the state and local level to guide instructional and assessment priorities, but were also included as part of the supervision and professional development process. In addition, some states adopted instructional frameworks as performance

scoring rubrics that determined teachers' evaluation of their effectiveness. We identified and included these terms in this category*: instructional frameworks or models (varied titled models), student outcomes, student learning objectives, student learning goals, instructional standards, learning standards, and performance standards.*

Ultimately, we identified and categorized these intersecting terms according to their unique purposes, differentiated interpretation, and application at the state- and local levels. This level of analysis provides a nuanced analysis of the data within and across states, representing the potential conflation or overlapping of supervision and evaluation terms and practices.

(4) Data trend analysis and display of findings.

Finally, we organized the data into comparative data tables to analyze and quantify state- and local control TSES dynamics and trends across the whole 50 state data set. We organized and cross-referenced our final categories of TSES policy development and implementation at the state- and local-levels: (1) requirements placed on TSES models, (2) requirements on TSES models based on student outcomes, (3) requirements of TSES models on student outcomes and teacher ratings, (4) embedded TSES development and student growth, (5) use of formative feedback in summative evaluation and TSES development, and (6) use of self-reflection in TSES model.

We then transferred the tabular data and codes to create six corresponding graphical map displays, shown and discussed in the findings section as Figures 1-6. These graphical displays provide theoretically grounded representations of six major patterns that we identified within and across the states

as a national sample of TSES policy development and implementation at the state- and local-levels.

Findings

Drawing on our 50-state analysis, we describe how state-level TSES plans are governing local districts' development and implementation of TSES frameworks. We organized our descriptive analysis and findings into six subsections, aligned with Figures 1-6: 1) Application of required external frameworks in TSES model development, 2) Student outcome requirements on TSES models, 3) Requirements for student outcome measures included in teacher ratings, 4) Requirements of student growth included in TSES development, 5) TSES models' formative feedback required with summative evaluation, and 6) Use of teachers' self-reflection required as an evaluative data point in TSES models.

Our analysis shows how TSES continue to be heavily regulated by federal- and state-level policy makers, constraining the control and authority of local systems leaders to develop and implement TSES that are grounded in the power of their respective communities.

Application of required external frameworks in TSES model development

Our first set of findings indicate that a majority of U.S. states governments have designed policies that preference the use of externally developed instructional frameworks to guide local TSES development. As shown in Figure 1, our analysis demonstrates the limited capacity for local districts' control of TSES models, with just 26% (13 states) allowing LEAs to select their own TSES model using an established state framework. This leaves a significant majority of states (74%) requiring specific TSES model development, with 52% of these requiring the use of predominantly

Danielson or Marzano models (26 states), and 22% requiring implementation of other identified state-level model (11 states). The local effects of external mandates have resulted in districts and states selection and reliance on external tools and instructional models designed to standardize instructional supervision and evaluation processes.

This finding suggests the challenges that systems-level leaders to try to balance, or merge, external influences and requirements meant to improve instruction with locally developed priorities and pedagogical innovations.

We acknowledge that within school districts that teachers are likely empowered to apply the external frameworks in different ways, and to select and emphasize various elements of the frameworks to improve instruction and student outcomes aligned with local TSES initiatives. However, the limitation of LEAs to select their own TSES model highlights how much control has been lost to state policy requirements, particularly over the course of NCLB and ESSA implementation.

This policy-driven model places school systems in the position of focusing most on adapting and adhering to external mandates, and working to ensure that external mandates are supporting, and not constraining local system initiatives.

This approach also assumes that one-size-fits-all models presume to identify and address teachers' instructional and students' learning needs at the local level. This potentially stifles local instructional, assessment, and professional development innovations guided by meaningful TSES systems that are mutually developed and implemented in collaboration by local administration, teacher leaders, and teachers.

Figure 1

Application of Required External Frameworks in TSES Model Development
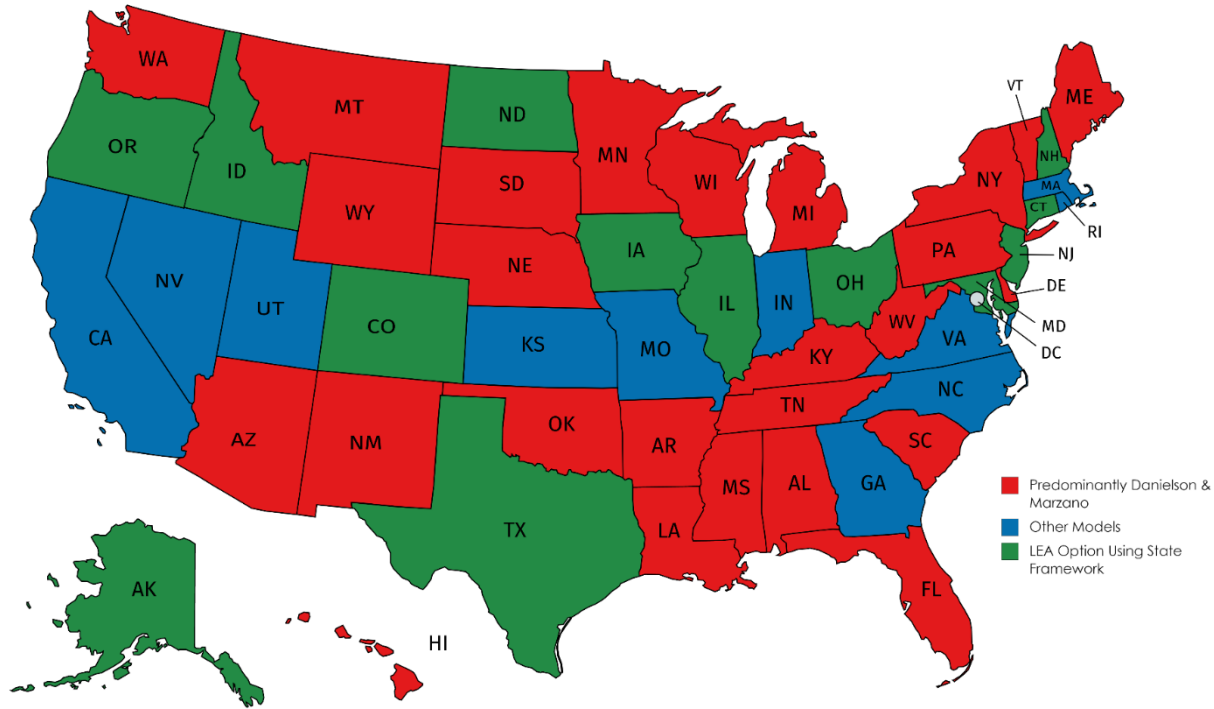
Student outcome requirements on TSES models

Integrated over time as part of NCLB and ESSA policies, states have continued to require student outcome evidence as part of local schools and teachers' progress. In our second set of findings, we separated out student outcomes from student outcome measures (in Figure 4) because states provide different definitions and applications of the term *student outcomes*. Based on the data, we also separated out local implementation of assessment tools to inform instruction and professional development, versus varied types of quantifiable outcome measures on standardized tests intended to determine

district, school, or teacher effectiveness over time. Local school systems' control to establish how and in what ways TSES models account for student-based outcomes are largely determined by state-level policies.

Figure 2 shows that roughly ¼ of all states (26%) remain neutral on placing requirements on TSES models to include student outcomes in the evaluation of teachers (13 states). Just 14% of states suggest, but do not require, the use of student outcomes (seven states), while 60% require the use of student outcomes in their TSES models (30 states). The smallest minority of the states that neither have requirements on TSES model

development, nor require student outcomes be used in TSES models. These four states, Iowa, Maryland, New Jersey, and Oregon, permit the maximized flexibility and development of TSES and local control of teacher development and evaluation regarding student outcomes.

This finding indicates that systems level leaders should collaboratively create a blend of local and state assessments that represent community learning priorities and teachers' respective content areas of practice. There is ample evidence of continued focus on

student learning among states and districts under ESSA; however, there exists a variable amount of national emphasis on the policy-practice connections between student learning outcomes and teacher supervision and evaluation systems. These distinctions become more evident when we drilled down into the states that indicate requirements for student outcomes, where we wanted to understand how the states were using student outcomes related to teachers' professional practices and evaluation ratings. We describe this part of our analysis in the next section.

Figure 2

Student Outcome Requirements on TSES Models
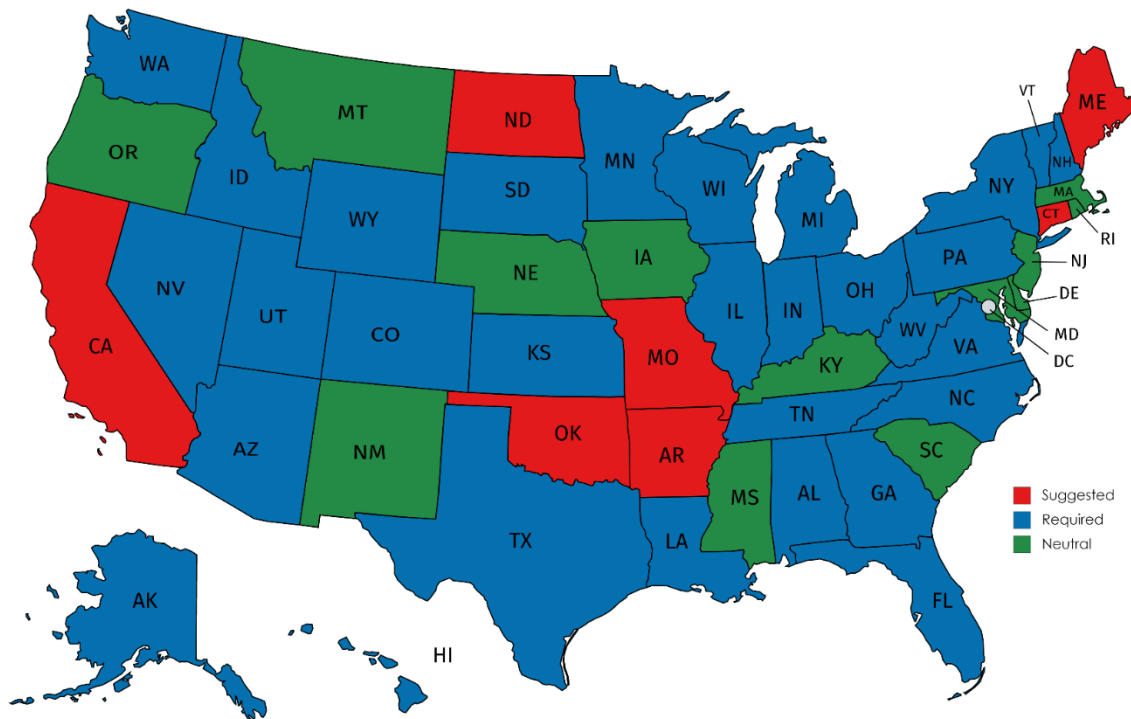
Requirements for student outcome measures included in teacher ratings

Given the national split regarding the requirements of student outcomes in TSES models, we extended our analysis to determine student outcome designations and how the states used this outcome measures data. In our third set of findings, we found that among the states that required student outcome measures, the states applied these measures in part to determine local teachers' performance ratings.

Figure 3 shows the requirements placed on TSES models as well as the respective percentages applied to teacher ratings, with 10% (five states) applying 50% or more of student outcome measures to teacher ratings, 18% (nine states) applying 35% – 49% of student outcome measures to teacher ratings, and 12% (six states) applying 20% – 34% of student outcomes to teacher ratings. One other state (Indiana) requires 25% – 50% of student outcomes be tied to teacher ratings.

This leaves almost 60% (29 states) where the states the permit flexibility to local districts to determine how student outcomes are tied to teacher ratings. From among these 29 states, 10 states require student outcomes be tied to teacher ratings but do not define the amount for LEAs (20% of all states), while 19 states suggest or remain neutral (but do not require) student outcomes be tied to teacher ratings (38% of all states).

This finding indicates that systems level leaders have some flexibility in approximately half the country to determine how to apply student outcomes data to their respective district level TSES systems. However, a significant number of systems leaders need to develop local systems that account for broadly disseminated and administered state-level assessment data.

The evidence indicates that a majority of states' TSES assessment or growth model policies throughout the U.S., not local education agencies, determine how teachers and instructional leaders are assessed, impacting the professional support structures that can be put in place for educators at the local level. In previous iterations of contemporary federal level policy under NCLB and Race to the Top (RTTT), policymakers applied student learning in distinctive ways.

For example, under NCLB student outcomes were included as measurable student test scores in at least English Language Arts and Mathematics, and measurable ratings of performance were applied at the district and school level. During RTTT, states were required to connect student learning outcomes, in part, as a measured test score or as a standard of learning progress artifact, as part of an individual teacher's performance evaluation.

The persistent remnants of these NCLB and RTTT policy levers applied under ESSA indicate that when given flexibility, states differ regarding how student outcomes connect with and are material evidence of teacher growth or professional performance.

Figure 3

Requirements for Student Outcome Measures Included in Teacher Ratings
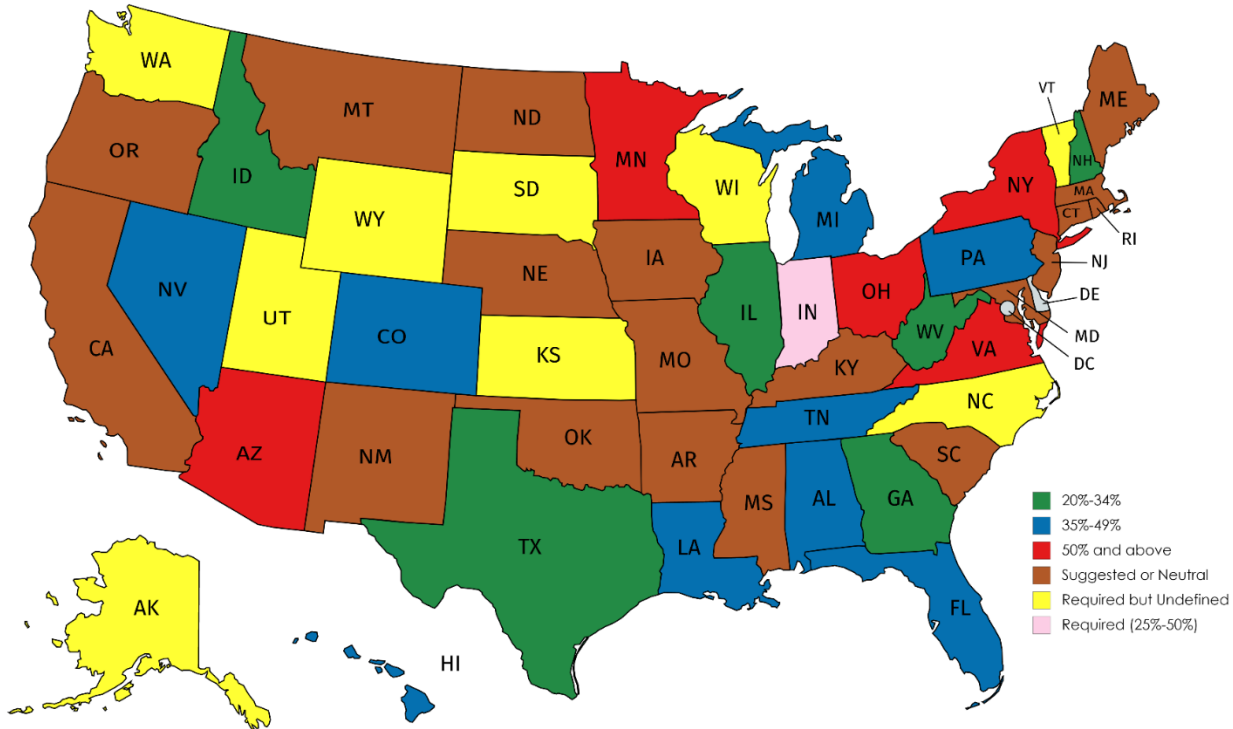
Requirements of student growth included in TSES development

Further developing our analysis of states' application of student outcome data, we discovered that states have continued to develop the use of student growth measures and models to determine local school effectiveness over time.

Figure 4 shows that 88% of all states have some sort of state requirements to show student growth and embed this in TSES model development. Of these, 28% (14 states) require the use of student learning objectives, 12% (6 states) require student growth models, such as value-added models or other state approved models, and 14% (7 states) require some sort

of student growth as measured through scores, percentiles, or other measures. Additionally, 34% (17 states) of all states do allow for student growth to be determined at the local level or through a combination of measures, while only 12% (6 states) of all states do not define any student growth requirement in their TSES models.

This finding suggests that systems-level leaders need to develop longitudinal assessment plans matched with developing multi-year instructional improvement plans, elevating the importance of strategic planning and collaborative visioning. The evidence points to a majority of states which have retained the top-down model of state-level

control to determine TSES model implementation during ESSA. States have continued to apply student outcomes broadly as a tool for instructional decision making, as means to determine elements of teachers' performance effectiveness, and as a tool to help determine district or school performance and student learning growth over time.

The ESSA era has not dissuaded a majority of state level policymakers to design and implement teacher effectiveness requirements that include student measures, and within this group of states, few states designate LEA's with the decision-making authority to determine contextualized applications of state codes. The evidence demonstrates that a persistent cultural shift has occurred post-NCLB and RTTT, and states have adopted a heavily bureaucratic, top-down stance towards evaluating teachers in terms of standardized, measurable assessment outcomes and methods to determine student learning and growth over time.

Figure 4

Requirements of Student Growth Included in TSES Development
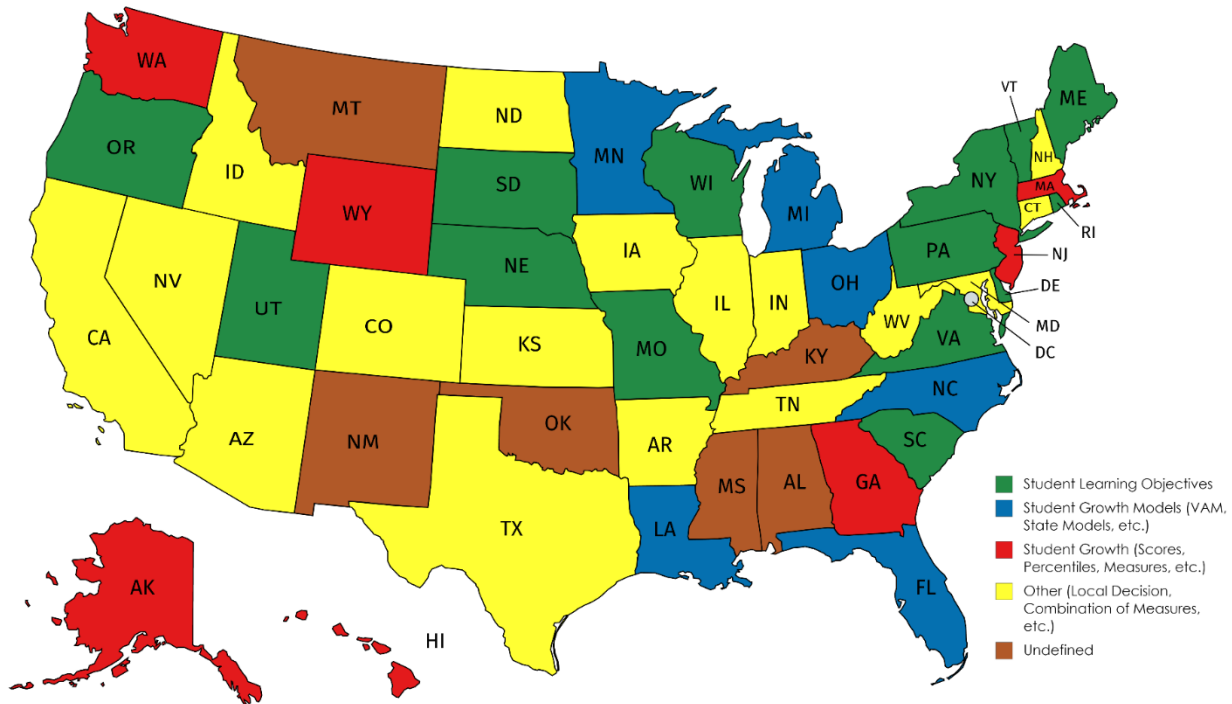
TSES Models' formative feedback required with summative evaluation

Focusing on the particular language of state-level TSES models aimed to support teacher growth and development, the final level of our analysis revealed that a significant majority of states codify the professional feedback processes as part of their respective TSES models. To determine how state TSES models structure the professional development of teachers and instructional leaders' practices, we analyzed the integration of formative feedback on teachers' instruction within TSES models.

Figure 5 displays that 82% (41 states) of all states embed formative feedback in their TSES models, which on the surface seems useful in the development of educators. However, the evidence revealed that this formative feedback is embedded in the summative evaluation of the TSES model, meaning that the ongoing formative feedback, which, by definition, is meant to be non-evaluative and help educators grow professionally, is actually used in the formal, summative evaluation of an educator. This means that just 18% (9 states) of state TSES models do not formally conflate the two respectively unique processes in state code.

This finding indicates that systems leaders in the majority of the country may need to establish and communicate purposeful structures and processes to separate the supervisory and evaluative processes that involve coaching, instructional feedback, and formative professional growth opportunities.

The conflation of supervision and evaluation potentially affects how teachers perceive and understand the dimensions of the supervision and evaluation process, impacting the nature of the administrative relationship which should be rooted in trust.

As we defined and described on our analytical framework, the formative functions of feedback provide teachers with elements of instructional coaching, peer feedback, professional learning, and support to enact ongoing pedagogical adjustments that teachers need to respond to students' needs. As an essential part of these processes, formative feedback requires a distinct separation from the aspects of TSES models that involve evaluative scoring, performance labels, and potentially coercive professional impacts which can impact a teacher's career.

Although we acknowledge that is difficult to separate the daily, temporal elements of an administrator's practice that can mentally integrate their observations from informal observations and formal evaluative observations, we question the merged codification of these processes at the state level. Teachers need professional space to engage in space professional conversations, which encourages risk-taking, the development of trust, and can spur innovation at the local level. Additionally, we found additional evidence of conflated formative professional growth processes regarding teachers' reflection embedded within TSES models, which we describe in the next part of our findings.

Figure 5

TSES Models' Formative Feedback Required with Summative Evaluation
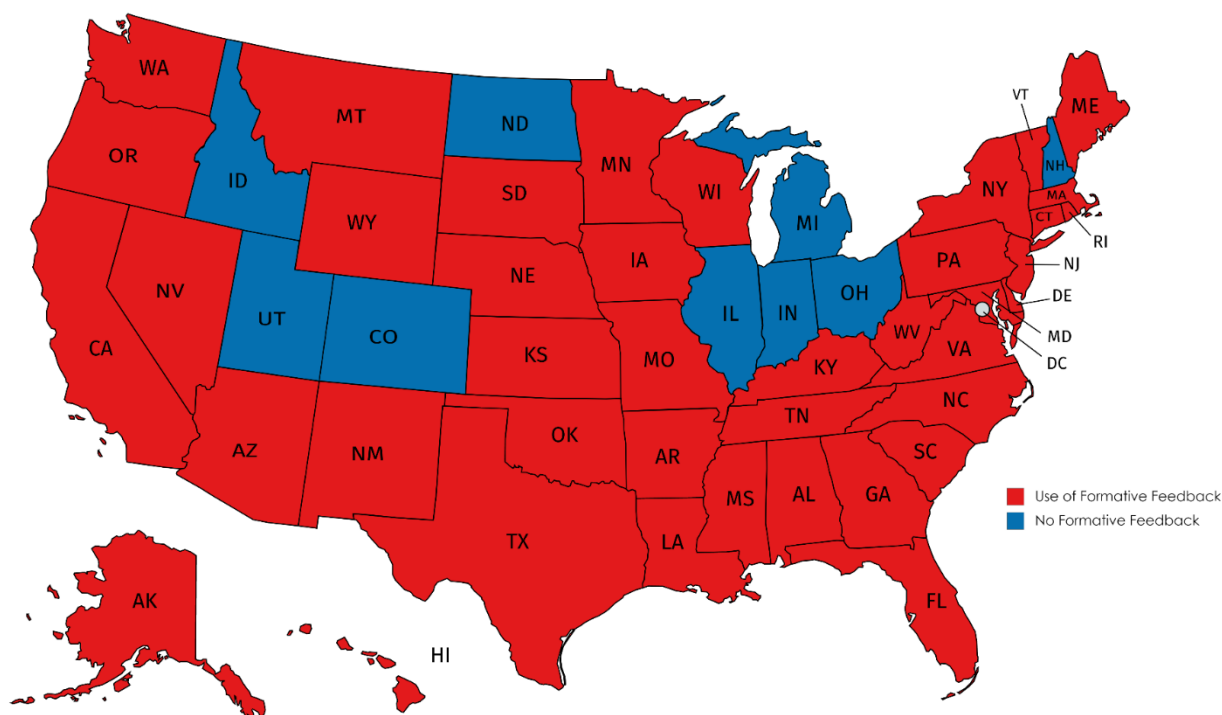
Use of teachers' self-reflection required as data point in TSES models

We discovered that the formative, and intrinsic process of teachers' reflective practices were integrated into TSES models at the state level but were in some state contexts were conflated with evaluative functions.

Figure 6 shows 78% (39 states) of all TSES models require the use of self-reflection as an aspect of summative evaluation, even though self-reflection is an internal process that an evaluator cannot actively see or observe.

An evaluative-focused reflective process would require teacher-principal dialogue, paired with subjective, external criteria to judge a teacher's reflective practices. This diverges from the intent of reflective practices, which encourages teacher reflection as a means to adjust instruction, innovate methods, and experiment to continuously refine their practices. As such, just 14% (7 states) use self-reflection solely for professional growth and/or professional development, 6% (three states) do not define the use of self-reflection in any way, and one state (Florida) varies the use of self-reflection based on LEA determination.

This finding can potentially impact on how district level leaders consider how to develop and implement structures and processes that support the development of teachers' and instructional leaders'

engagement and professional relationships within TSES systems. Although the process of reflection is often part of external instructional frameworks and models, the process has been overwritten by evaluative structures and policies that are using reflection in ways that are potentially counterproductive to teacher growth and development.

Given the diverse nature of the teacher workforce and positioning each professional as a lifelong learner with different styles, experiences, and identities, the codification of reflective processes or modalities limits the expressive nature of personal and professional reflection, and thereby limits its potential to

impact pedagogical development. Matched with formative feedback, the overlapping, evaluative emphasis which is codified in a significant number of state contexts potentially short-circuits the intended and research-based practices of reflection, coaching, and informal professional supports necessary to support teachers' growth and development during their careers. Systemically at the state and district levels, these conflated formative and summative processes can also impact the professional aspects of collaboration, trust, and safety necessary to develop positive professional cultures focused on instructional improvement and growth in student learning.

Figure 6

Use of Teachers' Self-Reflection Required as Evaluative Data Point in TSES Models

Discussion and Implications for Systems-level Leadership Practice

The purpose of this study was to investigate how 50 U.S. states have developed their respective TSES policies to provide local guidance to school district leaders and teachers. To review, we posed two research questions: 1) During ESSA implementation, what state-level TSES policy requirements govern local school districts' TSES development and implementation; and 2) What are the implications of these state-level TSES requirements for systems-level leadership practices at the district level?

In response to question one, our findings indicate that a majority of U.S. states have developed and implemented potentially prescriptive requirements that determine how local school districts support and evaluate leaders' and teachers' professional growth and performance. The evidence demonstrates that these policies rely on several elements that govern local district practices, namely external instructional frameworks, measures of student outcomes and growth over time, and mandated application of formative feedback and reflective practices that are simultaneously integrated into summative teacher evaluation processes.

These requirements potentially conflate instructional supervision practices that are intended to support teachers' growth and development and restrict local innovation and experimentation to improve student learning and experiences in schools. The findings also highlight the infiltration and influence of the educational improvement industry and profit driven economy that is potentially a factor in the development of standardized TSES policies at the state-level. Evaluation approaches drawn from products in the marketplace, such as applying score rubrics connected to

instructional frameworks, were never intended to become metrics for teacher performance and evaluation. This potentially creates a complex problem caused by an overreliance on tools that are unreliable methods to gauge teacher performance, compounded by evaluations based on student data, which may be interpreted against external frameworks or state assessments that are misaligned and being misapplied.

The evidence suggests that states and schools continue to grapple with how to provide adequate support for teachers and leaders at the local levels, relying on state-level mandates to structure leaders' and teachers' improvement efforts. In a significant majority of states, policymakers' default response has become a top-down accountability approach which indicates continued efforts to apply standardized methods to teachers' professional development, leaders' instructional supervision, and formalized teacher evaluation practices.

Despite the best research-based teaching and learning practices across the student grade-levels and content areas, policymakers continue to reduce instruction to a series of standardized list of classroom cues or teacher behaviors, neglecting the complexities of holistic pedagogy and student-centered instruction embedded within community contexts.

The emphasis on the standardization and accountability of TSES processes present significant challenges to systems-level leaders at the district-level, who have obligations to cultivate collaborative, responsive, and contextually relevant practices with the stakeholders in their respective communities. As a result of our findings, there are several

implications for systems-level leadership practice.

In response to research question two, we interpreted the evidence from our findings to discuss three main implications for systems-level leaders' practices: 1) prioritize localized visions for teaching and learning, 2) clarify teacher supervision and evaluation processes, and 3) provide instructional leaders with TSES professional development. Our goal is to provide practical suggestions to help leaders develop and implement locally responsive TSES systems within the ESSA policy era.

Systems leaders need to prioritize localized visions for teaching and learning

Citing evidence of policymakers' emphasis on top-down state-level TSES mandates, systems leaders will need to develop and implement contextually relevant visions for teaching and learning practices with teachers, building level principals, and community members.

Based on our analysis, a majority of state policymakers are requiring districts' use of externally developed teaching frameworks to determine local instructional and assessment practices. These external frameworks are designed to be generic, content-neutral, and vacant of contextual relevance. If systems-level leaders seek to prioritize community-oriented goals and employ pedagogical practices that are responsive to teachers' and students' needs, then leaders will have to enact proactive steps to integrate top-down mandates.

These steps will require a learner-centered approach to ensure teachers and leaders are working in unison and engaged in meaningful instructional supervision and evaluation processes. For example, teachers

and leaders will need to collaboratively create professional development that defines unique grade-level, content-area, and assessment applications that can be merged with external instructional frameworks. These applications necessitate the development of pedagogical language to bridge leader-teacher feedback conversations during the supervision and evaluation process, to build shared, site-level understandings of pedagogy, content, and assessment across unique, diverse classroom contexts.

As designed, these external frameworks are not intended to effectively meet students' learning goals, or teachers' professional growth needs, so systems leaders will need to develop these connective processes locally to provide meaning to local stakeholders.

Systems leaders need to clarify teacher supervision and evaluation processes

Citing evidence that indicates the conflated application of teachers' formative, reflective, and growth-oriented processes within TSES state-level evaluation policies, systems leaders will need to develop purposefully compartmentalized domains of supervision and evaluation processes to sustain professional growth cultures that protect and encourage professional safety, trust, and innovation.

We recommend a clearly articulated plan which outlines in what domains these respective processes are going to be applied, where a broadly defined, collaborative effort integrates the use of peer coaching, mentoring, teacher leadership, and non-evaluative personnel positions to delineate supervision from evaluative areas of practice.

These local TSES systems will need to address two main areas of overlapping

supervision and evaluation procedural tensions that emerged from our data analysis.

(1) **Clarify purposes for assessment data in TSES models.**
First, systems leaders will need to continue to develop local assessment practices and cultures that support ongoing student and teacher learning. Our study indicated two overlapping assessment processes conceptualized through ESSA policies. One process applies student outcomes data, which was described as locally developed assessment tools, district assessment monitoring, or integration of state assessment data systems, which is part of TSES models. Another process applies state-level longitudinal student growth data modeling through the administration of large-scale state assessments, district-wide assessment measures, or other assessment tools linked to respective state TSES models. Among both, TSES models include assessment data are applied to both supervision and evaluation processes at the local district level, and a majority have adopted these assessment purposes for teacher evaluation. Systems leaders will need to define specific purposes for assessment within respective supervision and evaluation practices, particularly given that assessment data can be used as formative monitoring of student learning coupled with teacher innovation to respond to students' needs on a daily basis. It is through these formative, experimental activities that teachers can demonstrate the most adaptive practices to expand their repertoire and document evidence of increased student learning.

(2) **Engage in supervisory dialogue that supports teachers' reflection and formative growth.** Second, systems leaders will need to consider how they support and facilitate teachers' reflective practices. In practice, the lines have always been temporally blurred between supervision and evaluation, and we acknowledge that on behalf of the administrator, it is difficult to mentally separate these situated moments of practice. However, state-level ESSA policies have taken this a step further and formally codified the feedback and reflective processes as related practices that span supervision and evaluation interchangeably. State-level ESSA policies have codified teachers' reflection within TSES models, integrating mandated evaluation of reflective practices via the application of external teaching frameworks. The process of feedback, reflection, and pedagogical dialogue was never intended to be a codified, evaluated practice, and requires teachers to engage with peers and administrators through relationships built on trust, transparency, vulnerability, and safety. Systems leaders can compartmentalize and protect the formative, supervisory functions of feedback and reflective practice by designing peer-reliant procedures that utilize dialogue and reflection with peer coaches, teacher leaders, colleague walk-throughs, and teacher mentoring at the grade, content, and team levels. These local practices can selectively buffer and supplement how external teaching frameworks are applied in relevant ways, addressing the local districts' needs and priorities through collaborative engagement

which necessitates the development of reflective pedagogical cultures and capacity building at the organizational level.

Systems need to provide instructional leaders with TSES support

Related to our previous suggestions, but important to state independently, systems leaders need to consider how to provide their building-level leaders, teacher leaders, instructional coaches, and teacher mentors with ongoing TSES professional development and support.

Engaging these individuals and instructional leadership teams in planning and implementation conversations about the purposes of supervision and evaluation and communicating their respective roles to teachers is critical to developing trust and positive cultures focused on innovation and risk-taking. While policy guidance can be helpful, it is a minimum marker for practice, and is limited by generalized tools, forms, and frameworks that may be disconnected from local vision, strategic goals, and community priorities.

Conclusion

Our analysis indicates that during the ESSA era, state-level policymakers have continued to develop and implement standardized methods for teacher supervision and evaluation, raising the likelihood that local control of teacher development is a potentially outdated concept in educational governance practice.

State-level policymakers are increasingly becoming the main source of authority in the development and implementation of local TSES processes, even when provided with more federal options and flexibility to initiate greater levels of local district governance.

The emphasis on the standardization and accountability of TSES processes present significant challenges to systems-level leaders at the district-level, who have obligations to cultivate collaborative, responsive, and contextually relevant practices with the stakeholders in their respective communities.

In a country which is increasingly diverse, and where local communities comprised of students, families, and teachers require personalized supports, ESSA represents a paradoxical problem for systems leadership practice.

In response, systems leaders will need to collaboratively reaffirm shared community commitments and take actions to retain local control of teacher development and instructional priorities.

Author Biographies

Douglas Wieczorek is an associate professor at Iowa State University's Educational Leadership, Organizations, and Policy. A former K-12 and university administrator, he researches how principals and teachers interpret and respond to state and federal educational policy changes. E-mail: dwieczor@iastate.edu

Israel Aguilar is an assistant professor in the department of organization and school leadership at the University of Texas, Rio Grande Valley. He has experience as a former school administrator in Dallas, TX., and a teacher in Austin, TX., and in Brownsville, TX. He currently researches social justice issues in the context of K-12. E-mail: israel.aguilar@UTRGV.edu

Ian Mette is associate professor of educational leadership at the University of Maine. His research includes supervision and the democratization of schools. His book, *A New Leadership Guide for Democratizing Schools from the Inside Out*, is coauthored with Carl Glickman. E-mail: ian.mette@maine.edu

References

Author Three. (2017).

Bowen, G.A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27-40.

Close, K., Amrein-Beardsley, A., & Collins, C. (2018). State-level assessments and teacher evaluation systems after the passage of the every student succeeds act: Some steps in the right direction. Boulder, CO: National Education policy Center.

Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of state's teacher evaluation systems post-every student succeeds act. *Education Policy Analysis Archives, 28*(58), 1-31.

Debray-Pelot, E., & McGuinn, P. (2009). The new politics of education: Analyzing the federal education policy landscape in the post-NCLB era. *Educational Policy, 23*(1), 15-42.

Donaldson, M., Mavrogardato, M., Youngs, P., Dougherty, S., & Al Ganem, R. (2021). "Doing the real work": How superintendents' sensemaking shapes principal evaluation policies and practices in school districts. *AERA Open, 7*(1), 1-16.

Edgerton, A.K. (2019). The essence of ESSA: More control at the district level? *Phi Delta Kappan, 101*(2), 14-17.

Every Student Succeeds Act, 20 U.S.C. 6301 (2015).

Ford, T.G., & Hewitt, K. (2020). Better integrating summative and formative goals in the design of next generation teacher evaluation systems. *Education Policy Analysis Archives, 28*(63), 1-34.

Gagnon, D.J., Hall, E.L., & Marion, S. (2017). Teacher evaluation and local control in the US: An investigation into the degree of local control afforded to districts in defining evaluation procedures for teachers in non-tested subjects and grades. *Assessment in Education: Principles, Policy, & Practice, 24*(4), 489-505.

Hazi, H. (2018). Coming to understand the wicked problem of teacher evaluation. In Zepeda, S.J. & Ponticell, J.A. (Eds.), *The Wiley Handbook of Educational Supervision* (pp. 183-207). Hoboken, NJ: John Wiley & Sons.

Kim, J., & Sun, M. (2021). The implementation and potential effects of teacher evaluation under local control. *School Effectiveness and School Improvement, 32*(2), 279-305.

Kirst, M.W. (1984). The changing balance in state and local power to control education. *Phi Delta Kappan, 66*(3), 189-191.

Koppich, J.E., & Esch, C. (2012). Grabbing the brass ring: Who shapes teacher policy?
  Educational Policy, 26(1), 79-95.

Lane, J.L. (2020). Maintaining the frame: Using frame analysis to explain teacher evaluation
  policy implementation. *American Educational Research Journal, 57*(1), 5-42.

McGhee, M.W. (2020). The supervision-evaluation debate meets the theory-to-practice
  conundrum: Contemplations of a practitioner turned professor. *Journal of Educational
  Supervision, 3*(2), 1-5.

Mcintyre, D.J., & McIntyre, C. (2020). The evolution of clinical practice and supervision in the
  United States. *Journal of Educational Supervision, 3*(1), 5-17.

No Child Left Behind of 2001, P.L. 107-110, 20 U.S.C. 6319. (2002).

Pauffler, N.A., King, K.M., & Zhu, P. (2020). Promoting professional growth in new teacher
  evaluation systems: Practitioners' lived experiences in changing policy contexts. *Studies
  in Educational Evaluation, 65*, 1-9.

Race to the Top Act of 2011. H.R. 1532- 112th Congress. (2014).

Schneider, J., & Saultz, A. (2020). Authority and control: The tension at the heart of
  standards-based accountability. *Harvard Educational Review, 90*(3), 419-445.

Silverman, D. (2000). Analyzing talk and text. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook
  of qualitative research* (2nd ed.), (pp. 821-834). Thousand Oaks, CA: Sage.

Stosich, E.L. (2020). Central office leadership for instructional improvement: Developing
  collaborative leadership among principals and instructional leadership team members.
  Teachers College Record, 122(9), 1-42.

Tuytens, M., Devos, G., & Vanblaere, B. (2020). An integral perspective on teacher evaluation:
  A review of empirical studies. *Educational Assessment, Evaluation, and Accountability,
  32*, 153-183.