

Teacher & Principal Evaluation Policy

A word about NCTQ's newest approach to tracking states' teacher policies

This report is the first in a series by the National Council on Teacher Quality (NCTQ) that examines the current status of states' teacher policies. Updated on a two-year cycle, each will cover a specific area of teacher policy. This report focuses on state teacher policies governing what states require in **evaluations of both teachers and principals**. The next edition will cover states' teacher preparation policies, and the third edition will cover states' compensation and personnel policies.

These reports are all drawn from data collected for NCTQ's [State Teacher Policy Database](#) (STPD), where we capture the full breadth of states' teacher policies, updated on a two-year cycle. They serve the important purpose of highlighting key trends across the nation on teacher policies, showing where the 50 states and the District of Columbia stand relative to one another.

In addition to the findings and trends presented here, users can access more detail on our database, such as states' individual scores and our analysis behind each score. Our database includes recommendations for each state tailored to its needs. We also invite states to offer their own commentary that we then publish. Historical information on policy areas, many of which NCTQ has been tracking for more than a decade, continue to be available in prior editions of NCTQ's *State Teacher Policy Yearbooks* (2007-2017) and *State of the States* reports (2012, 2013, and 2015), all of which are available to the public.

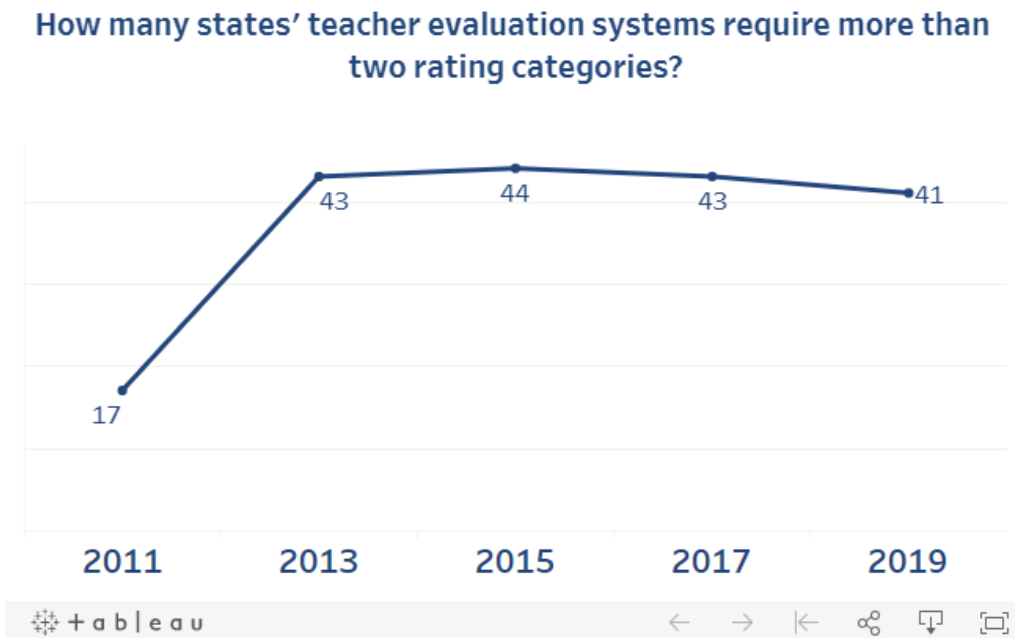
This report and the two that follow replace NCTQ's long-running *State Teacher Policy Yearbook*. We offer this new format in response to user requests for shorter, more digestible guides.

NCTQ is grateful to state education agencies for their gracious cooperation in our work, both recently and over the past dozen years. These partnerships have been critical in helping to ensure the accuracy of this final product.

Background

Historically, many teacher and principal evaluation systems have failed to yield meaningful, actionable data,^[1] leading most states (as well as school districts) to reform their educator evaluation systems over the past decade.^[2] These reforms, including increasing the number of possible ratings teachers and principals could earn, supplementing the measures on which educators are evaluated, and increasing the frequency and impact of evaluations, were all designed to make evaluation systems more accurately reflect individual educators' strengths and weaknesses. They also aimed to better distinguish the full range of educator talent.^[3]

Figure 1.

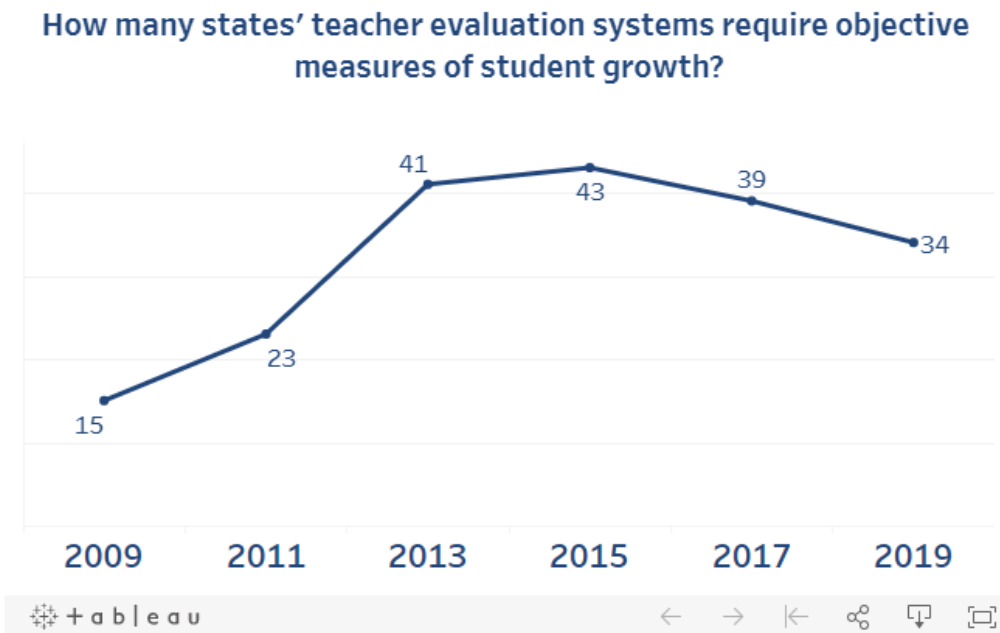


States' adoption of these new policies was remarkably swift. For example, in 2011, only 17 states maintained teacher evaluation systems that had more than two possible ratings (e.g., "satisfactory" and "unsatisfactory") that a teacher could earn, meaning that most school systems could not formally differentiate between merely adequate and truly exceptional teacher performance.^[4] In a span of only four years, by 2015, this number skyrocketed to 44 states.^[5] In contrast to many of the findings presented here, this number remains relatively stable, with 41 states continuing to commit to systems with more than two rating categories today.

Additionally, most historical teacher evaluation systems relied exclusively on subjective data, primarily based on principal observations of their teachers. In 2009, only 15 states required

objective measures of student growth in teacher evaluations; by 2015 this number increased nearly threefold to 43 states.^[6]

Figure 2.



However, as swiftly as states moved to make these changes, many of them have made a hasty retreat.^[7]

Over the past four years, many states have made modifications to their evaluation systems that are poorly supported by research literature; some have even abandoned their new systems altogether.^[8] Among the 43 states that made substantive changes to their evaluation systems within the last decade, nearly two-thirds (30) made at least one modification that runs counter to the research-supported evaluation practices NCTQ tracks. Some made a wholesale retreat, notably the **District of Columbia** and **Kentucky**, with **Arkansas**, **Oklahoma**, and **Wyoming** not far behind.

States' reasons for making these policy changes are undoubtedly as varied as the states themselves. Nevertheless, the U.S. Congress's reauthorization of the *Elementary and Secondary Education Act* of 1965 (ESEA) as the *Every Student Succeeds Act* (ESSA) in 2015 marks a notable inflection point. ESSA's enactment signaled the end of a period of heightened federal activity that included two initiatives, Race to the Top and ESEA flexibility, both of which incentivized states to develop and implement more objective teacher and principal evaluation systems. In the absence of these incentives, much of the momentum behind adopting and implementing more rigorous educator evaluation systems ground to a halt.

Figure 3.

Activity in state educator evaluation policies, 2015 - 2019

State	Teacher Evaluation Requirements							Principal Evaluation Requirements				
	Student growth data	Student growth data based on the state standardized test	More than two ratings categories	Annual evaluations for all teachers	Requires or explicitly allows survey data	Multiple observations	Support for struggling teachers via improvement plans	Student growth data	Annual evaluations for all principals	Requires or explicitly allows survey data	Annual observations for all principals	Support for struggling principals via improvement plans
Alabama	Green	Grey	Grey	Red	Green	Grey	Grey	Grey	Green	Grey	Green	Green
Alaska	Red	Red	Green	Grey	Green	Grey	Green	Red	Green	Green	Green	Green
Arizona	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Grey
Arkansas	Red	Grey	Green	Red	Green	Grey	Red	Red	Grey	Green	Grey	Red
California	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Red	Grey
Colorado	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green
Connecticut	Green	Green	Green	Green	Green	Red	Green	Green	Green	Green	Green	Green
Delaware	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Grey	Green
District of Columbia	Red	Grey	Red	Green	Red	Red	Grey	Red	Red	Red	Grey	Grey
Florida	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Grey	Green
Georgia	Green	Green	Green	Green	Red	Green	Green	Green	Green	Green	Green	Red
Hawaii	Green	Red	Green	Grey	Green	Grey	Green	Green	Red	Green	Green	Green
Idaho	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Red	Grey
Illinois	Green	Green	Green	Grey	Grey	Grey	Green	Green	Grey	Red	Green	Grey
Indiana	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green	Green
Iowa	Grey	Grey	Red	Grey	Green	Grey	Green	Green	Green	Green	Green	Green
Kansas	Red	Red	Green	Green	Green	Green	Grey	Red	Grey	Green	Green	Green
Kentucky	Red	Red	Green	Green	Green	Grey	Red	Red	Red	Red	Red	Grey
Louisiana	Green	Green	Green	Green	Grey	Grey	Green	Green	Green	Grey	Green	Green
Maine	Red	Red	Green	Grey	Green	Green	Green	Green	Grey	Green	Green	Grey
Maryland	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Massachusetts	Green	Red	Green	Grey	Green	Green	Green	Green	Green	Green	Green	Green
Michigan	Green	Green	Green	Grey	Green	Grey	Green	Green	Grey	Green	Grey	Green
Minnesota	Green	Green	Green	Green	Green	Grey	Green	Green	Green	Green	Green	Green
Mississippi	Green	Red	Green	Green	Green	Green	Red	Red	Green	Green	Green	Grey
Missouri	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Montana	Grey	Grey	Grey	Grey	Green	Green	Grey	Grey	Grey	Grey	Grey	Grey
Nebraska	Grey	Grey	Green	Green	Red	Green	Green	Grey	Grey	Green	Green	Green
Nevada	Green	Grey	Green	Red	Green	Green	Grey	Green	Red	Green	Green	Green
New Hampshire	Grey	Grey	Green	Green	Green	Green	Green	Green	Green	Green	Green	Grey
New Jersey	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
New Mexico	Red	Red	Green	Green	Green	Green	Green	Green	Red	Green	Green	Green
New York	Green	Red	Green	Green	Grey	Green	Green	Green	Green	Green	Green	Green
North Carolina	Red	Red	Green	Green	Green	Grey	Green	Red	Green	Green	Green	Grey

Green Maintained or implemented policy since 2015 Red Retreated since 2015 Grey No progress prior to or since 2015

State	Teacher Evaluation Requirements							Principal Evaluation Requirements				
	Student growth data	Student growth data based on the state standardized test	More than two ratings categories	Annual evaluations for all teachers	Requires or explicitly allows survey data	Multiple observations	Support for struggling teachers via improvement plans	Student growth data	Annual evaluations for all principals	Requires or explicitly allows survey data	Annual observations for all principals	Support for struggling principals via improvement plans
North Dakota												
Ohio												
Oklahoma												
Oregon												
Pennsylvania												
Rhode Island												
South Carolina												
South Dakota												
Tennessee												
Texas												
Utah												
Vermont												
Virginia												
Washington												
West Virginia												
Wisconsin												
Wyoming												
Number of states that have retreated since 2015	11	12	3	5	5	2	4	10	6	6	5	4
Number of states with no progress prior to or since 2015	6	13	7	24	15	39	14	7	15	14	18	23
Number of states that made progress prior to 2015	32	25	41	22	29	9	31	33	29	26	22	23
Number of states that have made forward progress since 2015	2	1	0	0	2	1	2	1	1	5	6	1

■ Maintained or implemented policy since 2015
 ■ Retreated since 2015
 ■ No progress prior to or since 2015

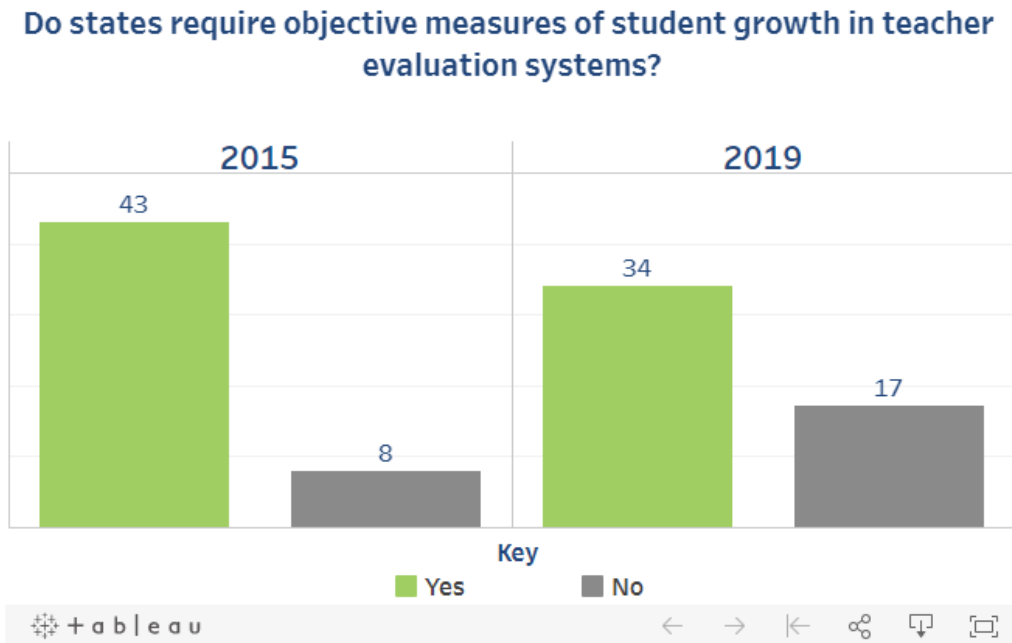
This figure depicts changes in key state teacher and principal evaluation policies between 2015 and 2019. It demonstrates that, in general, more states have retreated from research-backed policies over the past four years than have adopted them.

For more, see our full **State of the States 2019: Teacher and Principal Evaluation Policy** report, available at: <https://www.nctq.org/pages/State-of-the-States-2019:-Teacher-and-Principal-Evaluation-Policy>

Teacher Evaluation Policy

Objective Measures of Student Growth

Figure 4.

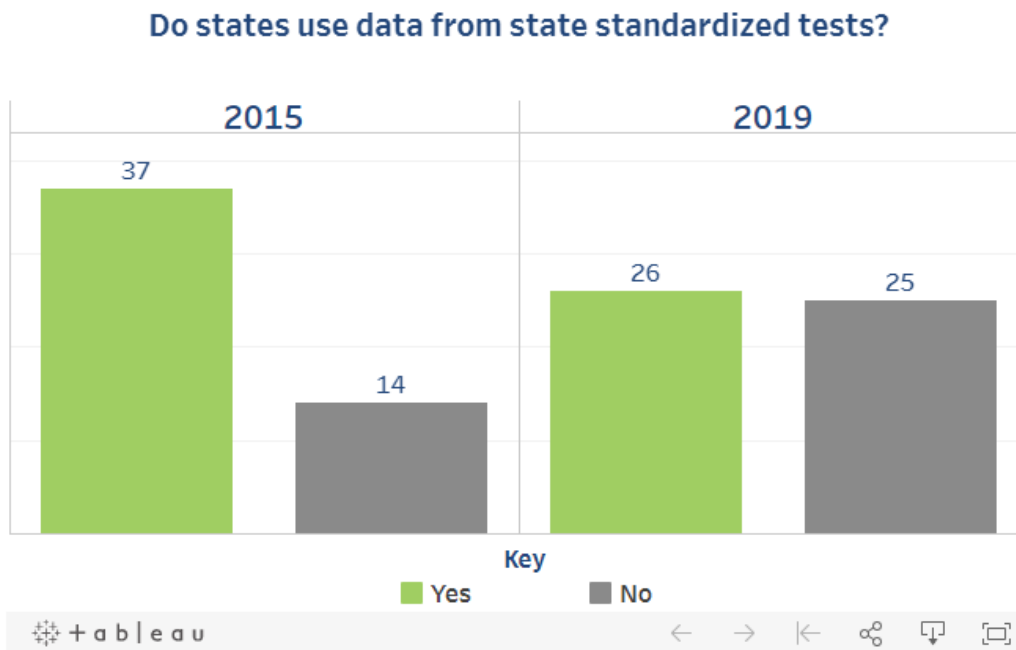


Formal teacher evaluations are more likely to be a fair measure of teacher performance when based on multiple measures.^[9] They are also more likely to be a valid measure of performance if they include objective measures, given the well-documented limitations of using only subjective measures, such as classroom observations.^[10] From 2011 to 2015, states worked quickly to incorporate objective measures of student growth into evaluation systems, responding to research, policy incentives, and the paramount importance of considering teachers' contributions to student learning when evaluating their performance in the classroom.^[11]

Currently, 34 states require teacher evaluations to include objective measures of student growth, down from a high of 43 in 2015. While 10 states (**Alaska, Arkansas, Kansas, Kentucky, Maine, New Mexico, North Carolina, Oklahoma, Wisconsin, and Wyoming**), as well as the **District of Columbia**, have dropped the requirement that teacher evaluation systems include objective measures of student growth since 2015, two states (**Alabama** and **Texas**) have added a student growth requirement during this same time period, for a net reduction of nine (eight states and the **District of Columbia**).

The Role of State Tests to Measure of Student Growth

Figure 5.

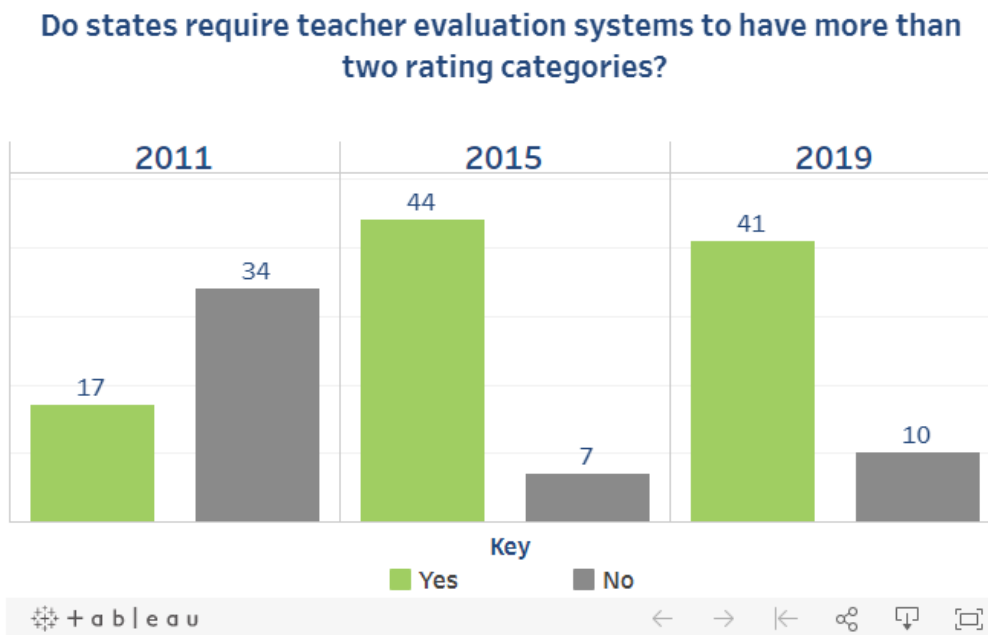


Even among the 34 states that continue to require some objective measure of student growth, fewer of those states require their state test to be the source of such data. Approximately one-quarter (8) of the 34 states that require teachers to be evaluated at least in part on objective measures of student growth do not currently require the state's standardized test to be the source of those data for at least some teachers (e.g., teachers of tested grades and subjects).

States undoubtedly have myriad reasons for making this change, including political shifts in some states and implementation challenges in others. By eliminating the state test as the required data source for calculating growth measures, states provide their districts with more say in *how* to measure their teachers' impact on student learning. Districts' use of measures such as district assessments, student portfolios, and student learning objectives to determine all teachers' contributions to student growth may help to build more buy-in for evaluation systems from educators. On the other hand, this shift means that states can no longer reliably compare teacher performance among districts. It also likely requires more monitoring and oversight on the part of the state to ensure that districts preserve the objectivity of their systems.

Evaluation System Rating Categories

Figure 6.



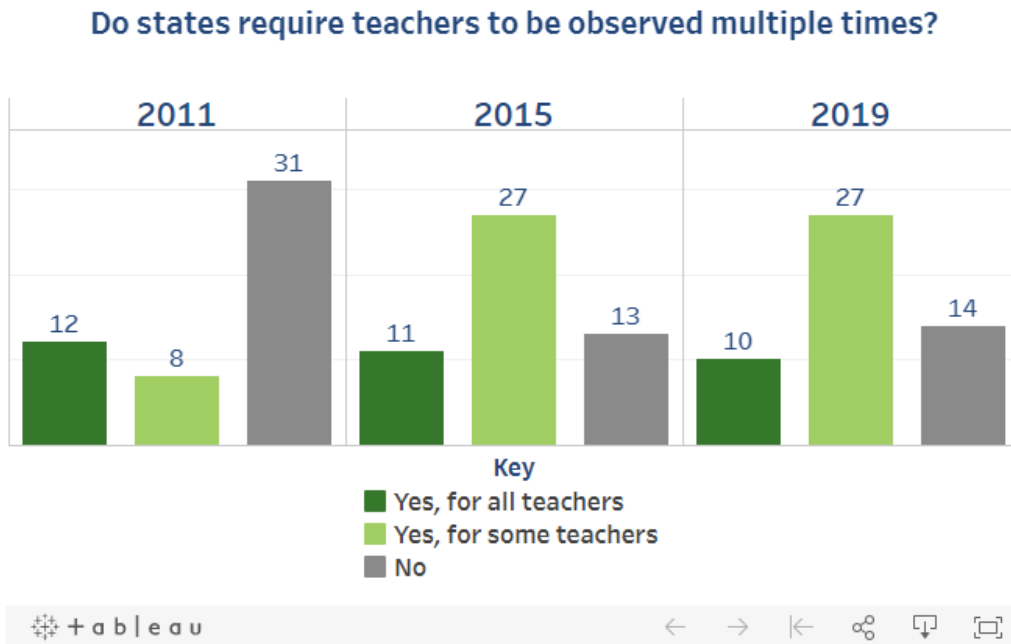
Historically, states and districts have struggled to formally differentiate between teachers making different contributions to students' learning and lives. Under evaluation systems with only two rating categories, almost all teachers earned the same rating of satisfactory (or its equivalent).^[12]

When this problem surfaced, many states moved to add additional rating categories, in an effort to ensure that their evaluation systems provided more nuanced information. Most states remain committed to this principle. The number of states maintaining more than two ratings has remained relatively stable since 2015.

As of 2019, more than 80 percent of states (41) require an evaluation with at least three rating categories. Only nine states (**Alabama, California, Iowa, Montana, Nebraska, New Hampshire, Vermont, Wisconsin, and Wyoming**) and the **District of Columbia** currently adhere to a binary system.

Observation Measures

Figure 7.

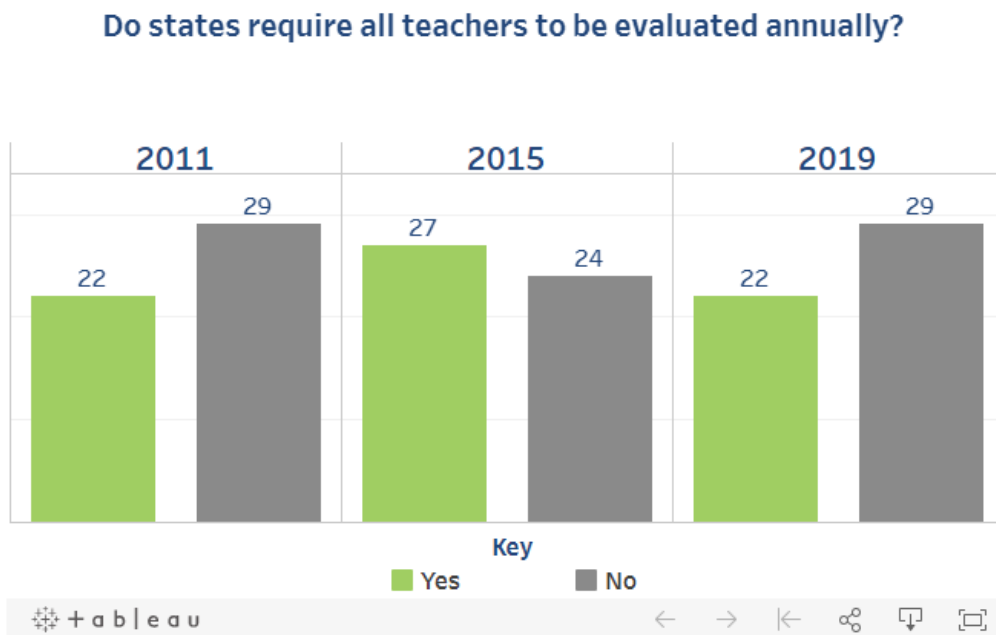


Observations by a school leader, administrator, or third-party evaluator continue to play a prominent role in teacher evaluations. While no state has eliminated classroom observations within the last decade, a number of states have modified their approach to this critical component. Many states reduced the weight of observations in a teacher's overall rating, made modifications to achieve more reliability, and pressed for early and frequent observations of new teachers.

One step many states took before 2015 was to require multiple observations of teachers, as studies have found that more than one observation is necessary to accurately capture a teacher's performance.^[13] Across the country, most states have not backed away from requiring multiple observations of at least some teachers. One fewer state requires multiple observations for all teachers in 2019 as did in 2015.

Evaluation Frequency

Figure 8.

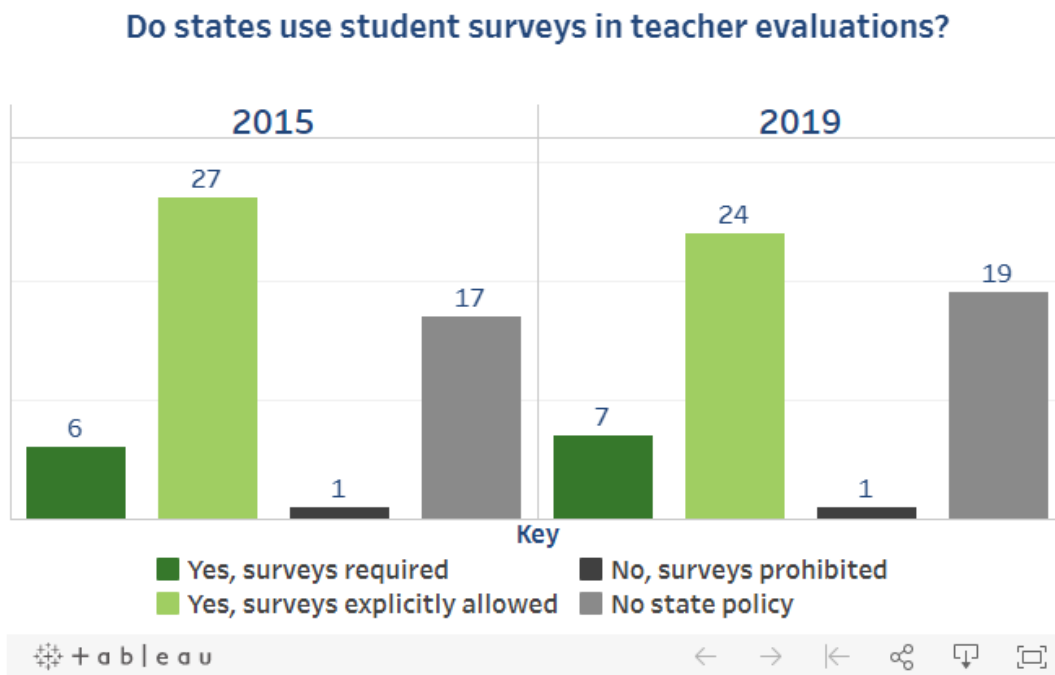


Over the past four years, there has been some movement in how frequently states require teacher evaluations. The years in between 2011 and 2015 represented a tipping point in states' requirements regarding evaluation frequency. In 2011, fewer than half of all states (22) required annual, summative feedback of all teachers. By 2015, more than half of all states (27) maintained this requirement, mirroring the practice across other professions.

As of 2019, only 22 states maintain this requirement, representing a complete reversion to the status quo in 2011.

The Role of Student Surveys in State Evaluation Systems

Figure 9.



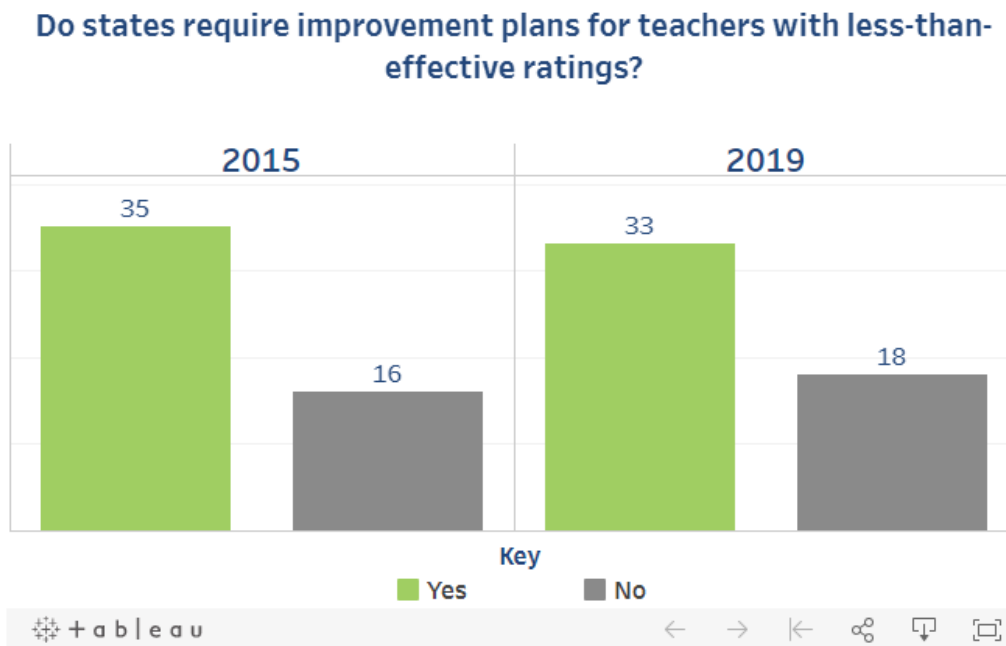
Surveys of a teacher's students, when well designed, can function as a meaningful component of teacher evaluation systems. A well-constructed student feedback survey correlates with student learning gains, providing schools with another independent source of teacher performance, alongside state tests.^[14] When included as part of a teacher's summative evaluation rating, student surveys contribute to teacher evaluations that are more reliable and valid than evaluations that rely solely on classroom observations by an administrator.^[15]

These results should not be altogether surprising because student surveys are based on tens of thousands of hours of experience with a teacher (e.g., 25 students, six hours a day, 180 days a year), versus a handful of hours by an external observer.

Despite the research-backed benefits of surveys, slightly fewer states in 2019 (31) either require or explicitly allow districts to factor in student survey data than in 2015, when 33 states did so. Only one state, **New York**, explicitly prohibits the use of student survey data in teacher evaluations.

Using Evaluation to Drive Targeted Support

Figure 10.



Before 2015, states had also worked to attach more consequences or mandated actions for teachers who were struggling, including actions designed to ensure that professional development is targeted to individual teacher needs.

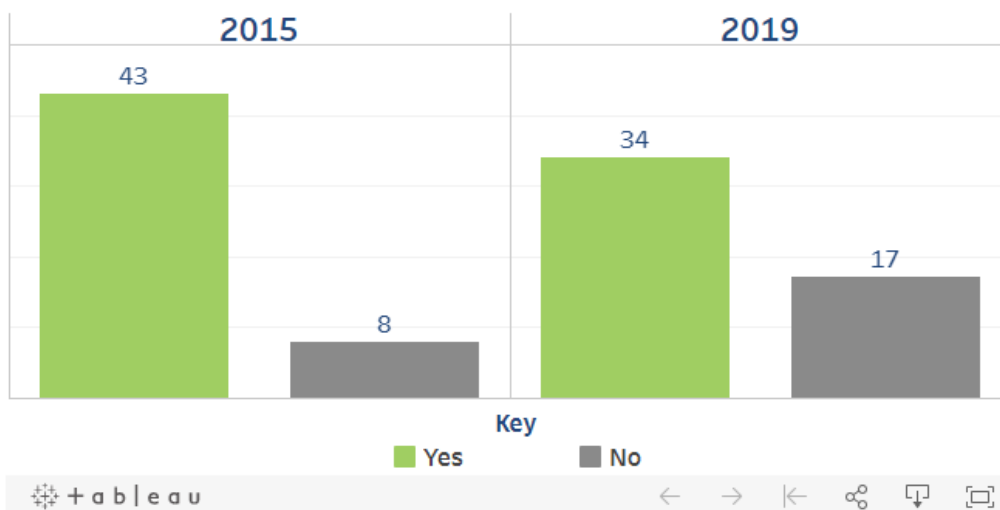
While four states (**Arkansas**, **Kentucky**, **Mississippi**, and **Oregon**) have withdrawn the requirement that teachers identified as the most in need of support receive targeted intervention by way of an improvement plan, two states (**Idaho** and **Iowa**) have added this requirement for a net reduction of two states between 2015 and 2019. Currently, more than two-thirds of states (33) continue to require that teachers who earn the lowest ratings are placed on improvement plans.

Principal Evaluation Policy

Objective Measures of Student Growth

Figure 11.

Do states require objective measures of student growth in principal evaluations?

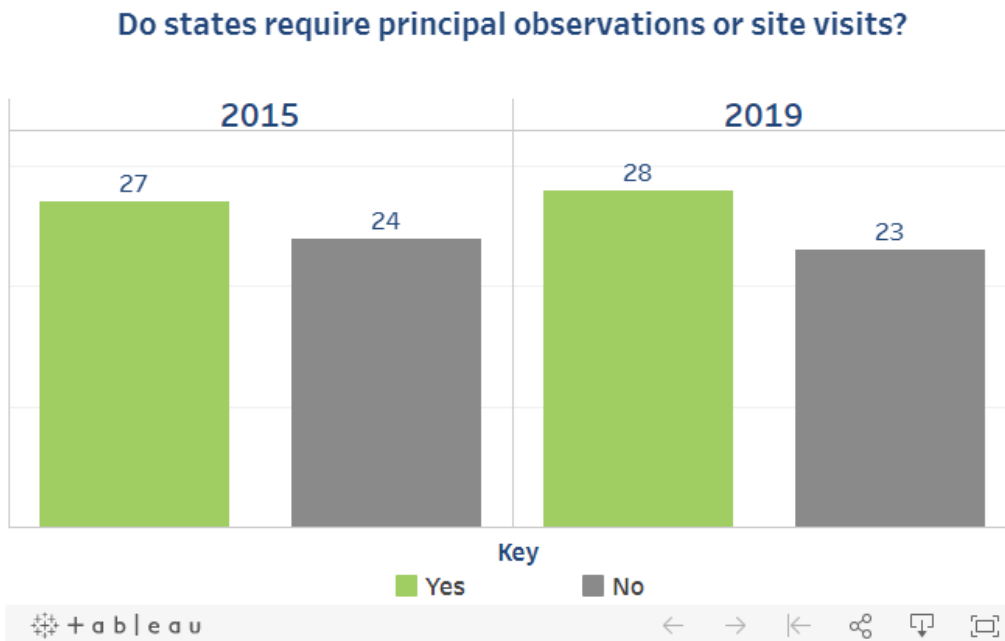


Great principals have a clear impact on the most important in-school factor influencing student learning and lives: teachers. Effective principals are more adept at managing their teacher faculty and are more successful at retaining effective teachers and removing ineffective ones.^[16]

Since 2015, the pattern of policy changes in principal evaluation has largely mirrored that of teacher evaluation policy. Many states have steadily weakened their principal evaluation system requirements. Currently, as compared to the high point of 2015, eight fewer states and the District of Columbia require all principals to be evaluated based, at least in part, on objective measures of student growth. Yet, principal quality continues to vary and significantly affects student achievement and school climate.^[17]

Observation Measures

Figure 12.



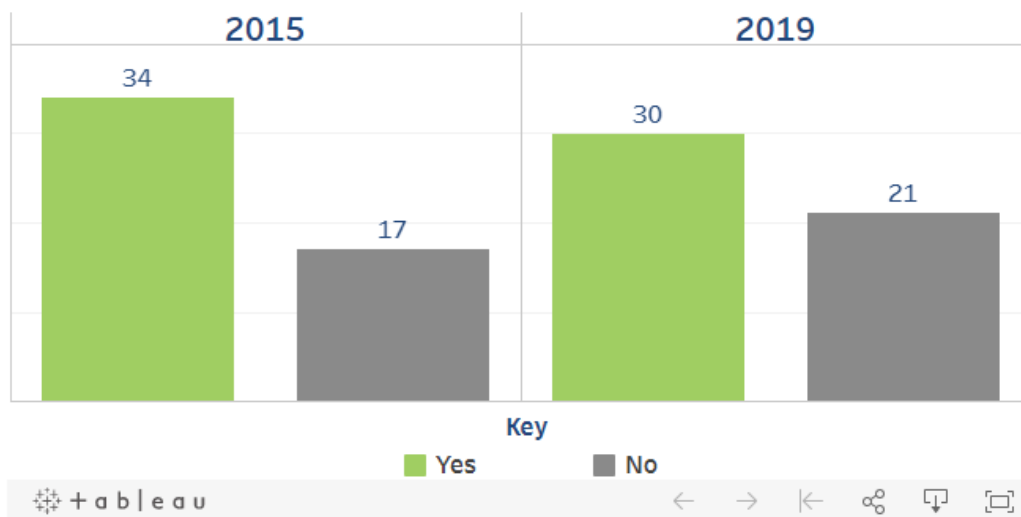
Similar to teacher evaluation systems, the strongest principal evaluation systems are comprised of multiple measures. Just as observing teachers gives valuable insight into their classroom practices and provides an opportunity to give targeted feedback, observations of principals by their supervisors can yield similarly rich information. Observations can be helpful in part because principals' effectiveness in organizational management and instructional planning is related to positive student and teacher outcomes.^[18]

There has been relatively little movement over the past four years in the number of states requiring principal observations or site visits. In 2015, 27 states explicitly required all principals to be observed annually or visited on site. In 2019, 28 states maintain this requirement.

Evaluation Frequency

Figure 13.

Do states require all principals to be evaluated annually?

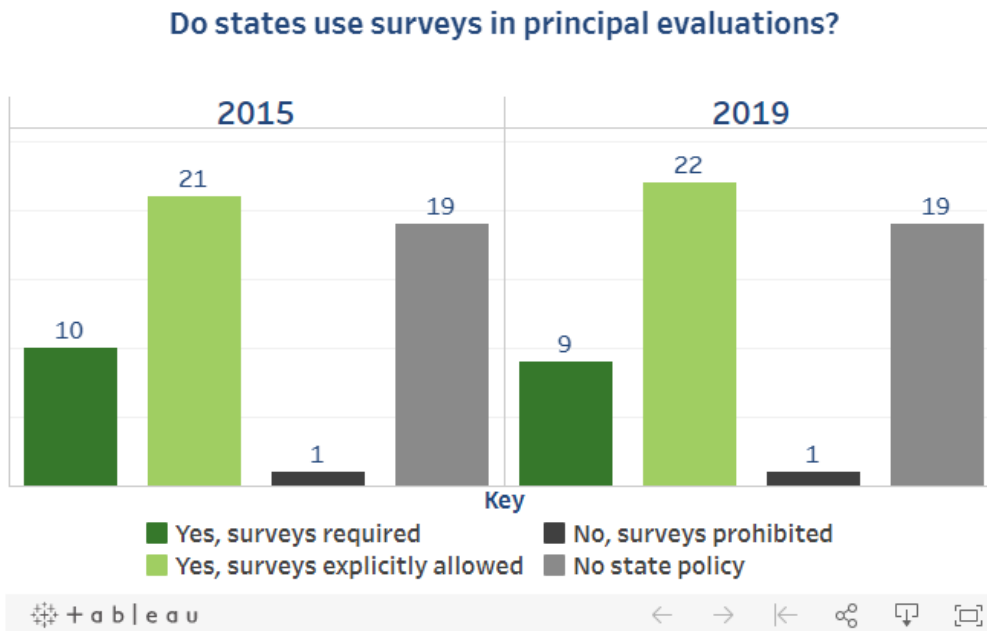


Failure to provide all principals, including exceptional ones, with annual feedback deprives them of the information necessary to improve their school leadership. Lack of annual feedback on principal quality also puts policymakers at a disadvantage as they seek to ensure that all students and teachers have access to effective school leaders.

Between 2015 and the present, there has been a decrease in the number of states requiring annual evaluations for all principals. In 2015, 34 states maintained this requirement; only 30 do so today.

The Role of Surveys in State Evaluation Systems

Figure 14.



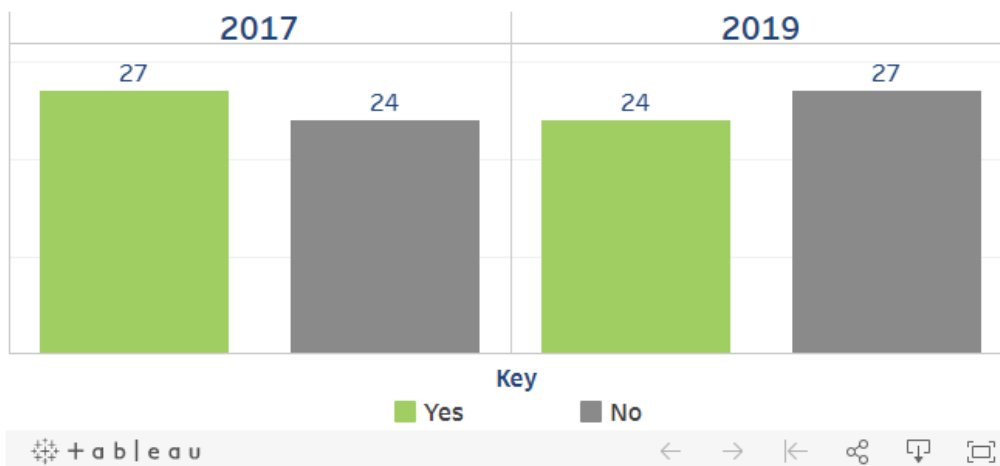
Survey data, whether from a school leader's students, teachers, or community, can provide a more comprehensive picture of a principal's performance than observation data alone. Given the significant variation in principals' effectiveness and the impact strong principals have on student achievement, in-school discipline, parents' perceptions of schools, and school climate,^[19] survey data can add important context to principal evaluations.

Since 2015, there has been little movement in the number of states requiring or explicitly allowing surveys to be a component of principal evaluation, with nearly two-thirds (31) of states using these tools in both 2015 and 2019. Just as in teacher evaluations, **New York** stands alone in prohibiting the consideration of surveys in principal evaluations.

Using Evaluation Systems to Drive Targeted Support

Figure 15.

Do states require improvement plans for principals with less-than-effective ratings?



Strong principal evaluation systems, like strong teacher evaluation systems, can help practitioners improve their practice. Principal evaluation systems that require targeted interventions for those most in need of support can help achieve necessary improvements. Requiring that principals who earn evaluation ratings of less-than-effective are placed on improvement plans helps to ensure that struggling school leaders receive the targeted support they need. Ultimately, stronger school leadership helps drive improved school outcomes.^[20]

Since NCTQ first collected data on this metric two years ago, the number of states that require targeted interventions through articulated improvement plans for principals earning less-than-effective ratings has declined. Currently, fewer than half of all states (24) maintain this requirement, as compared to more than half of all states (27) that did so in 2017.

Recommendations

Over the past decade, changes to educator evaluation systems have been undeniably new and different. Necessarily, they have required continuous modification and adjustment. However, some of these modifications have weakened and reduced available data, thereby decreasing the ability of policymakers and educational leaders to make meaningful personnel decisions. Unfortunately, it is hard to attribute many of these changes to anything other than a desire to revert to the status quo; that is, to former systems that generally failed to provide the information necessary for individual teachers to improve their practice and for policymakers to make strategic personnel decisions. Further, few of these changes were supported by best practice or research literature.

As states continue to monitor, iterate, and improve teacher and principal evaluation and support systems, they should be mindful of important research findings establishing core tenets to a strong evaluation system. These tenets are as follows:

- Objective measures of student growth substantially improve the validity of evaluations and help to capture the more genuine range of educator talent within a school, district or state.
 - Because neither the best ratio of objective to subjective data nor the optimal mix of data sources are firmly established, states should carefully monitor the components and results of their educator evaluation systems and make any necessary adjustments to the weights of different system components.
- Multiple observations of all teachers, including observations conducted by more than one individual, improve evaluation system reliability.
- More than two rating categories, as compared to binary evaluation systems, increase the useful information available to individual educators and policymakers.
 - Detailed data can support policymakers' efforts to ensure that practitioners have access to appropriate resources and supports and individual educators' efforts to improve their practice.
- Annual evaluations benefit all educators, regardless of effectiveness levels.
- Survey data provide important information about an educators' performance.

As the country contends with increasing inequality, ensuring equitable access to effective teachers and school leaders remains paramount. Continuing to invest in and improve upon the systems that provide information about educator effectiveness is essential to ensure that all students, particularly vulnerable students, have equitable access to effective educators, and that practitioners have access to the necessary information to improve their practice.

Endnotes and References

1. Weisberg, D., et al. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: TNTP. Retrieved June 27, 2019, from <https://tntp.org/publications/view/the-widget-effe...>
2. Ross, E., et al. (2017). *2017 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on September 19, 2018, from https://www.nctq.org/dmsView/NCTQ_2017_State_Teac...
3. See *Components of a strong teacher evaluation system*. https://www.nctq.org/dmsView/Components_of_a_strong_eval_system
4. Jacobs, S., et al. (2011). *2011 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on June 27, 2019, from <https://www.nctq.org/publications/2011-State-Teach...>
5. Jacobs, S., et al. (2015). *2015 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on June 27, 2019, from <https://www.nctq.org/publications/2015-State-Teach...>
6. Jacobs, S. et al. (2009). *2009 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on June 27, 2019, from <https://www.nctq.org/publications/2009-State-Teach...> and Jacobs, S., et al. (2015). *2015 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on June 27, 2019, from <https://www.nctq.org/publications/2015-State-Teach...>
7. Ross, E., et al. (2017). *2017 State Teacher Policy Yearbook National Summary*. Washington, DC: National Council on Teacher Quality. Retrieved on September 19, 2018, from https://www.nctq.org/dmsView/NCTQ_2017_State_Teac...
8. See Figure 3.
9. Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, 46(3), 587-613; Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *The American Economic Review*, 102(7), 3628-3651.
10. Id.
11. Hanushek, E. A., & Hoxby, C. M. (2005). Developing value-added measures for teachers and schools. *Reforming Education in Arkansas*, 99, 104.; Clotfelter, C., & Ladd, H. F. (1996). Recognizing and rewarding success in public schools. In H. Ladd (Ed.), *Holding schools accountable: Performance based reform in education* (pp. 23-64). Washington, DC: Brookings Institution Press; Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service.

12. Weisberg, D., et al. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. New York, NY: TNTP. Retrieved June 27, 2019, from <https://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness>
13. Glass, G. V. (1974). A review of three methods determining teacher effectiveness. In H. J. Walberg (Ed.), *Evaluating Educational Performance* (pp. 11-32). Beverly Hills, CA: Sage; Travers, R. M. W. (1981). Criteria of good teaching. In J. Millman (Ed.), *Handbook of Teacher Evaluation* (pp. 14-22). Beverly Hills, CA: Sage; Xu, S., & Sinclair, R. L. (2002). *Improving teacher evaluation for increasing student learning*. Paper presented at the annual meeting of the AERA, New Orleans, LA.
14. Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9; Wagner, W., Gollner, R., Helmke, A., Trautwein, U., & Ludtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction, 28*, 1-11; Kane, T. J., & Cantrell, S. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle, WA: The Bill & Melinda Gates Foundation.
15. Wallace, T. L., Kelcey, B., & Ruzek, E. (2016). What can student perception surveys tell us about teaching? Empirically testing the underlying structure of the Tripod student perception survey. *American Educational Research Journal, 53*(6), 1834-1868.
16. Beteille, T., Kalogrides, D., & Loeb, S. (2009). *Effective schools: Managing the recruitment, development, and retention of high-quality teachers* (Working Paper 37). National Center for Analysis of Longitudinal Data in Education Research.
17. Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). *Estimating the effect of leaders on public sector productivity: The case of school principals* (No. w17803). National Bureau of Economic Research; Louis, K. S., Leithwood, K., Wahlstrom, K. L., Anderson, S. E., Michlin, M., & Mascall, B. (2010). Learning from leadership: Investigating the links to improved student learning. *Center for Applied Research and Educational Improvement/University of Minnesota and Ontario Institute for Studies in Education/University of Toronto, 42*, 50; Clark, D., Martorell, P., & Rockoff, J. (2009). *School principals and school performance* (No. w17803). National Bureau of Economic Research; Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning: A review of research for the Learning from Leadership Project*. New York, NY: The Wallace Foundation.
18. Grissom, J. A., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal, 48*(5), 1091-1123; Horng, E. L., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education, 116*(4), 491-523; Catano, N., & Stronge, J. H. (2007). What do we

expect of school principals? Congruence between principal evaluation and performance standards. *International Journal of Leadership in Education*, 10(4), 379-399.

19. Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). *Estimating the effect of leaders on public sector productivity: The case of school principals* (No. w17803). National Bureau of Economic Research; Louis, K. S., Leithwood, K., Wahlstrom, K. L., Anderson, S. E., Michlin, M., & Mascall, B. (2010). Learning from leadership: Investigating the links to improved student learning. *Center for Applied Research and Educational Improvement/University of Minnesota and Ontario Institute for Studies in Education/University of Toronto*, 42, 50; Clark, D., Martorell, P., & Rockoff, J. (2009). *School principals and school performance* (No. w17803). Cambridge, MA: National Bureau of Economic Research; Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. L. (2004). *How leadership influences student learning: A review of research for the Learning from Leadership Project*. New York, NY: The Wallace Foundation.
20. Clifford, M., Hansen, U. J., & Wraight, S. (2014). *Practical guide to designing comprehensive principal evaluation systems: A tool to assist in the development of principal evaluation systems*. Washington, DC: Center on Great Teachers and Leaders; Rice, J. K. (2010). *Principal effectiveness and leadership in an era of accountability* (Brief 8). Washington, DC: National Center for Analysis of Longitudinal Data in Education Research; Glasman, N. S., & Heck, R. H. (1992). The changing leadership role of the principal: Implications for principal assessment. *Peabody Journal of Education*, 68(1), 5-24.

To see a full review of each state’s teacher policies, visit: www.nctq.org/yearbook

Suggested Citation:

Ross, E. & Walsh, K. (2019). *State of the States 2019: Teacher and Principal Evaluation Policy*. Washington, DC: National Council on Teacher Quality.

Special thanks to Kelli Lakis and Lisa Staesina for providing research support to this project.