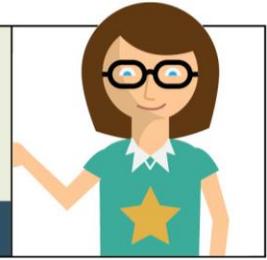


**EVALUATION, ACCOUNTABILITY, AND PROFESSIONAL DEVELOPMENT IN AN OPPORTUNITY CULTURE**

**A PRACTICAL GUIDE**

**STEP 3: EVALUATION CONTENT FOR NEW ROLES**



**EVALUATION REDESIGN ACTION STEPS**

All of these actions are taken by the **Evaluation Design Team, Human Resources, and/or OC Initiative leaders**. Principals take action for specific teachers in their schools annually, as indicated.

Use the *District Evaluation Content Plan* for all steps.

- Prepare to design evaluation content
  - Review input about evaluation content gathered from stakeholders and design team
  - Review *Lessons from Teacher Evaluation in Traditionally Staffed Schools*
- Define/select the job responsibilities for each role using role descriptions. (*Principals define for individuals annually*). Use Opportunity Culture job descriptions as a start (see the *Teacher and Staff Selection Toolkit*)
- Identify the behavioral competencies—habits of thinking and action—that are important for success in each role. Use Opportunity Culture suggested competencies for each role as a start (see the *Teacher and Staff Selection Toolkit*)
- Identify outcome measures and tools
  - Student learning outcomes, including clear, useful student growth measures
  - Conditions for learning, such as attendance and behavior/discipline
  - Instructional practice improvement
  - Others
- Determine ratings that define effective and highly effective on each measure
- Determine weighting of each measure in that evaluation category
- Determine if teachers will have one overall evaluation rating, and if so what the weighting of each category of measures will be
- Determine the amount of consistency required in performance (will teachers new to a role have a grace period or other “pressure valve” to allow ramp-up?)
- Choose, adapt, or create tools and assessments to measure/rate performance on job responsibilities, competencies, and outcome measures. (*Principals adapt to individual roles, where needed.*) Use Opportunity Culture **evaluation tools, job descriptions, and competencies** as starting resources for annual review materials.
- Match teachers to students, subjects, and teams at school level:
  - Matching teachers to the right students
  - Matching teachers to the right subject(s) for each student
  - Matching team leaders to other teachers they lead and develop, and
  - Matching team leaders to team teachers’ success and improvement
- Determine if district will calculate magnitude of reach for each teacher, and how
- Match formal accountability systems at district and state levels to roles. (*Principals inform and collaborate with Human Resources and OC Initiative leaders for matching.*) Use *Evaluation and Accountability Database Tool* to help.
- Record decisions on the *District Evaluation Content Plan*
- Use strategic, frequent communication to inform all affected

## CONSIDERATIONS AND GUIDANCE

This step covers critical design issues, organized into these sections for discussion:

- \* General Considerations
  - Rating and weighting
  - Performance consistency
- \* Specific Content Design Issues
  - Job responsibilities
  - Competencies
  - Outcomes: Student learning, conditions for learning, and instructional practice improvement
  - Matching to the right students, subjects, and teams at school level
  - Magnitude of impact, or “reach”
  - Matching formal accountability and data systems at district and state levels to roles
- \* Special Issues for Common Opportunity Culture Roles

### General Considerations

Evaluation content changes in an Opportunity Culture, because the roles require new behaviors, and interaction with more students and other teachers. Some content remains the same, but districts can use the redesign process to fix problems in existing evaluation systems.

This section discusses critical evaluation content decisions that every district must make when scaling up Opportunity Culture roles into many schools. It also includes model tools. As you determine what your district’s evaluation will cover, and how those elements will be measured and weighted, check your choices against *Lessons from Teacher Evaluation in Traditionally Staffed Schools* to ensure that your evaluation content reflects lessons from the past about what to include and what to avoid.

The Opportunity Culture Evaluation Design Team must decide whether to use or adapt the materials provided here; in all cases the team will still have critical design decisions to make—such as the weighting of different measures, and ensuring that the evaluation system aligns with career advancement requirements.

The Opportunity Culture initiative leader and team can present options to the Opportunity Culture Evaluation Design Team, or make decisions based

on that team’s prior input. The Opportunity Culture Evaluation Design Team should be kept informed of all decisions and have the chance to review and comment upon draft materials.

The best evaluation systems across sectors consider performance in job responsibilities, important outcomes, and the competencies—or habits of action and thinking—that allow a person to fulfill those responsibilities, achieve those outcomes, and succeed in a given role.<sup>1</sup> In an Opportunity Culture, magnitude of impact—or the number of students reached—is also critical, because it may determine the teacher’s position within career and pay paths, and because it communicates the importance of a teacher’s total impact.

In addition to guidance about each of these, this step contains special discussion of evaluation content for major Opportunity Culture roles and model evaluation content and tools. Districts and schools can adopt these tools, adapt them, or use the tools as one source of input. In all cases, the specific responsibilities that may be unique to a given person’s role should be added, and the specific students and subjects covered should match outcome data used in each evaluation.

Some critical design decisions span these areas: the weighting of each measure in an overall evaluation score, the ratings levels reflecting differing levels of performance on each measure, and the consistency required for performance, both ongoing and when a person is new in a role.

**Rating and weighting:** Districts (and/or states) must decide what **rating** on each measure and overall constitutes “effective” and higher levels of contribution, and how much each measure will affect the summary rating in that category and overall ratings (“**weighting**”).

In some states and districts, ratings by colleagues will merely inform principals’ final, overall rating of each teacher. In others, the ratings and weightings will be more formulaic—and in these cases, getting them right is especially important.

In all cases, try to design an evaluation system that includes annual improvement in measures, rating levels, and weighting, as well as the process steps to effect this improvement. Try not to adopt a system that will be static for multiple years—it will

inevitably be inferior to systems that a team of teachers and administrators can keep improving.

*Most important: The criteria for entering and remaining in advanced roles on the career paths should align with the measures, weightings, sub-scores, and rating levels in the evaluation system.*

The performance levels required to enter and remain in advanced roles should be levels that actually predict outcomes of students affected by each role.

For now, districts can make educated guesses about rating cut points using whatever data are available or common sense—ideally with *substantial* educator input for common-sense judgements.

The best districts will track ratings data over time and correlate these data back with student outcomes to determine *which other measures* and what *levels of performance* best predict, and thus support, student learning in different roles. More predictive measures should have increasing weight as districts improve their evaluation systems and also be used for selection and advancement opportunities, where possible. Research conducted across the Opportunity Culture initiative may reveal the best measures to weigh heavily, and what overall levels of performance correlate with success in advanced roles.

Districts may or may not have one **overall “score”** for teachers. If *not*, then the district must decide which measurement categories are the ones that will guide advancement opportunities for each role. If districts calculate one overall evaluation score, then *properly weighting the categories of measurement is essential*: When a district uses ratings to extend the reach of teachers with certain qualities, they will get far more of that quality and its consequences in schools. If, for example, competencies are weighted too heavily, schools may find that gaps emerge—job responsibilities undone and student outcomes falling short. The best systems will likely weight student outcomes significantly, invest in improving measures of student outcomes, and keep a significant weight on competencies and job responsibilities, both of which help teachers achieve student outcomes.

Districts also must add any other categories of measurement to those included here—for example, pre-existing measures used in teacher evaluations

that the district wishes to keep for all teachers—and include those in the weighting system.

Although it’s tempting to simply glue OC evaluation onto existing evaluation, districts scaling up OC roles into many schools should think about how to simplify, not simply add. Most evaluation systems already have too much detail to be valuable to teachers.

Use this chance to streamline and focus on the measures most helpful to teachers and predictive of success in new roles.

In model materials here, we are agnostic about the relative weighting of various measures, except that student growth should account for at least a third of overall teacher evaluation ratings. Student growth measures, though, need to be transparent—i.e., teachers need to understand how they are calculated and what they mean, ideally in months or years of student growth—and fair, reflecting continued learning about typical student growth trajectories. In addition, teachers cannot be held accountable for a full school year’s growth when students enter midyear.

We also recommend continuing to invest in improved student assessments that reflect applied learning—not just rote skills and memorization—and higher-order thinking skills.

On page 7 we provide an example of weighting with categories of performance typically used in an Opportunity Culture. This example illustrates how important it is to focus on measures with the most potential impact on student learning and teacher improvement, and only on ones relevant to the role. It also depicts a level of complexity that causes some states and districts to have each supervisor—in most cases the principal—choose final ratings in each category, rather than a complex weighting system.

In practice, some combination of *quantitative ratings*, outcomes and daily work by those who work most closely with a teacher, and *final*

---

***Streamline and focus on the measures most helpful to teachers and predictive of success in new roles.***

---

*judgement*, by the supervisor, on each category's final rating may be the path that many choose.

**Performance Consistency.** When determining eligibility for a teacher to obtain and keep an advanced role, districts may choose to give teachers some wiggle room to account for normal fluctuations in effectiveness. Even the best teachers have off years—due to student and school factors and personal matters—and this is particularly to be expected when starting a new role.

Accounting for normal fluctuations can be accomplished by requiring high-scoring evaluations only a certain number of years out of a total—such as two out of every three years, or three out of every four. For teachers new to roles, it can be accomplished with a yearlong “grace period” in which to learn a new job and begin performing at the expected level. Or both.

However, districts eager to see teachers perform quickly in new roles will not rely on grace-period learning alone: They will provide advance training specific to each role, peer-team co-planning and improvement, and coaching throughout the year to avoid performance dips for both new and experienced advanced-role teachers.

### Specific Content Design Issues

**Job responsibilities:** Leading evaluation systems for other professionals are not based on lists of tasks for jobholders to do. They include job *responsibilities*, which provide guidance but leave employees with some flexibility in how to achieve those responsibilities. For example, the model role descriptions for teachers provided on OpportunityCulture.org include these categories of responsibility, with more detail, but not specific tasks, for each:

- \* Planning and preparation
- \* Classroom environment
- \* Instruction
- \* Professional responsibilities

Roles requiring supervision and leadership have those duties woven into each of these four categories. For example, a multi-classroom leader's planning and preparation includes setting high expectations for the team, establishing methods

and materials that the team will use, setting direction to clarify the content and teaching process, and leading the team collaboration to plan and improve personalized instruction that is ambitious and includes higher-order thinking skills.

Role descriptions that are vague leave the school at risk of promoting less-ambitious teaching and learning. But role descriptions that list detailed tasks strip teachers of the responsibility and authority to think hard, try their best, innovate, and improve. Task lists can also be misleading for evaluating professionals and make it easier to fall into compliance mentality. For example, it is better for teaching-team members to note that a peer needs to improve the ambitiousness, higher-order thinking, and personalization/differentiation of lesson plans—and to give some evidence and examples—rather than to check “yes” on a task list that a teammate has completed paperwork required for lesson plans, however unambitious the plans are.

Suggested role descriptions with job responsibilities may be found in the **Teacher and Staff Selection Toolkit**.

### Competencies

Competencies are the habits of thought and behavior that predict success on the job. They distinguish people who on paper look the same—but in practice are very different in their ambition for students and ability to influence students and peers, among other qualities. Competencies are critical, because they help individuals fulfill responsibilities and achieve outcomes. Competencies also help provide a focal point for development. Rather than a leader merely telling a teacher “do x task,” the leader can help the teacher see the specific thinking and actions that will help the teacher “do x really well, consistently.”

Because advanced role opportunities on career paths are affected by competency evaluations, choosing the right competencies is very important. They should be ones that can not only rate teachers' effectiveness in *current* roles, but also indicate strengths that could be used in *advanced* roles and on teams, as well as weaknesses to develop.

Understanding strengths really well can help principals assemble strong teaching teams with complementary talents. Principals and multi-classroom leaders can also encourage advancement into roles in which a teacher is most likely to excel—roles that draw on strengths rather than magnifying weaknesses.

Suggested competencies for common roles can be found in the [Teacher and Staff Selection Toolkit](#).)

### Outcomes

Outcomes to consider are student learning, conditions for learning, and instructional practice improvement.

Student outcome measures for evaluation in an Opportunity Culture should incorporate the right measures and the right students, subjects, and other teachers for whom a teacher is responsible.

Outcomes measures should be:

- \* growth-focused, so that teachers are not penalized for taking students starting behind
- \* easy to understand, ideally translated into years or months of learning that help teachers plan and monitor student growth over the course of each year
- \* inclusive of critical thinking skills—analysis and applying concepts—not just memory and technical skills
- \* comparable across students in different classes and schools
- \* multi-measure, not only standardized tests, especially when single measures are especially unreliable or when single measures leave out important learning goals
- \* correlated with standardized test outcomes when measuring similar content, especially when no assessment that is comparable across schools and districts is available in a grade or subject
- \* inclusive of student work products that show both content knowledge and organizing skill, such as papers, reports, and projects—ideally assessed with *rubrics correlated* with standardized test outcomes
- \* inclusive of student learning, of course, but also conditions for learning, such as student attendance and disciplinary rates, and critical instructional practices. See the [Standard Evaluation Tools Package](#) for examples.

---

*In an Opportunity Culture, great teachers can help all teachers learn how to achieve high test scores not through excessive testing, but through high-engagement learning focused on the same academic content.*

---

Student outcomes—and by this we mean development in all areas, not just content knowledge—are the purpose of education. The validity and reliability of outcomes measurements must continue to improve to help teachers develop and to identify teachers who can best lead and develop peers on the job and serve more students very well. Perhaps this is why teachers in Opportunity Culture schools have embraced measuring student progress throughout the year.

But assessment of outcomes need not replace instruction that is engaging: student work that requires planning, conceptual engagement, and persistent thinking and action is compatible with outcome measurement. Scoring rubrics for this kind of work can be *correlated* with standardized tests. Students can then spend more time on valuable learning activities of varying sorts, including those that require making choices and producing an original end product, making learning more engaging overall.

Still, teacher-graded work is not the best source of reliable, wide-scale data about the impact of instruction on student growth. Standardized tests still have their place.

Great teachers already understand this. They know that successful test performance is a golden ticket to college and beyond, as well as to many of the better-paying technical and blue-collar jobs in our economy. In an Opportunity Culture, great teachers can help all teachers learn how to use high-engagement learning, not excessive test prep, to improve test scores. When scoring rubrics are correlated with standardized tests, teachers can track student growth during the year using a wide variety of learning activities.

However students are assessed, teachers will differ in their impact on student outcomes. Cross-sector research indicates that while outcome measures

differ by sector and role, the *distribution* of measured outcomes of professionals and managers is remarkably consistent—and nearly identical to the distribution of outcomes found in student outcomes among teachers.<sup>2</sup> Even if the measures of success change for teachers, the distribution—or the proportion of educators performing at the top, middle, and bottom on a measure—would likely remain stable. In addition, the variance in today’s student outcome assessments using value-added measures is no more than that found in performance measures commonly used in other professions.<sup>3</sup>

*Matching to the right students, subjects, and teams at school, district and state levels.* Outcome measures should match the specific students, subjects, and other teachers for which each teacher is actually responsible, and should reflect students’ growth and peer teacher improvement over a full school year. Growth measures must allow for shared attribution—holding multiple teachers accountable and giving them all credit for a student’s learning. This requires **annual planning** to make sure that each teacher’s scope of student responsibility is accurately reflected in the district and state evaluation system.

See “Special Issues for Common Opportunity Culture Roles” on page 9 for specific issues in matching teachers to the right students and

subjects, and for examples of how to assign accountability for teams and team leaders.

See the *Evaluation and Accountability Database Tool* for one way to organize each school’s roster to clarify each teacher’s and teacher-leader’s student and subject accountability. Districts can put this into formal accountability and human resource information systems.

- \* Principals can use this “roster of responsibility” created in the *Evaluation and Accountability Database Tool* to monitor the progress of the team, team leader, and individual teachers during the year—for example considering the midyear growth of each teacher-leader’s whole team of students.
- \* Districts should ensure that data in this tool is then communicated to those in the district and state who track formal evaluation outcomes and accountability.
- \* Districts must also ensure that interim student learning data collected—using digital instruction and assessment and other means—is reported to the right teachers, team leaders, and principals so that each person has the data needed to improve instruction and develop other staff during the year. This is essential for “Little e” improvement throughout the school year (for more on this, see [Step 2](#)).

### Special Issues: Transient Students, Students with Special Needs, and English Language Learners

**Transient students:** In addition to improving assessments overall, districts should clarify a policy that does not harm teachers when a student starts midyear, especially in schools with large transient student populations. Dealing with the specifics of this are beyond the scope of this guide, but districts should seek to align the interest of teachers and students who switch schools midyear. Such students need to achieve growth during partial years in multiple schools—especially students who are homeless or face other situations that contribute to learning losses. Teachers need fair evaluations that reflect only the time that a particular student was in their class. The best data systems will incorporate partial-year growth data into student growth-measurement systems, but do so fairly. The best districts will not treat this as a teacher evaluation issue, though. Instead, they will adopt policies to reduce student transience, such as providing bus service to economically disadvantaged students moving within the same district and collaborating with housing agencies to reduce moves.

**Other needs:** It is beyond the scope of this guide to provide detailed guidance on evaluation for students with special needs and English language learners. But districts must adjust evaluation of students and their teachers to encourage a focus on growth, motivating expectations for achievement, and proper attention to the social and emotional needs of learners in these populations.

### Example: Weighting Evaluation Measures and Categories

Here is a brief example of a way to weight evaluation measures using the categories of performance included in Opportunity Culture standard materials.

This is just one example and not a recommendation of weightings; each location must determine its own weightings in collaboration with teachers and administrators.

Measure weights are used to determine the score for each category of performance.

Category of Performance	Measure	Measure Weight in that Category	Category Weight in Overall Rating
<b>Fulfillment of Job Responsibilities</b>			25%
	<i>Planning &amp; Preparation</i>	.25 (or 25%)	
	<i>Classroom/School Environment</i>	.25	
	<i>Instruction</i>	.25	
	<i>Professional Responsibilities</i>	.25	
<b>Demonstrated Behavioral Competencies</b>			25%
	<i>Driving for Results</i>	.4 (or 40%)	
	<i>Influencing for Results</i>	.4	
	<i>Problem Solving</i>	.1	
	<i>Personal Effectiveness</i>	.1	
<b>Outcomes</b>			50%
	<i>Student Learning</i>	.75 (or 75%)	
	<i>Conditions for Learning</i>	.10	
	<i>Instructional Practice Improvement</i>	.15	

Each measure will typically have sub-measures. For example, each person can be rated on multiple competencies within each competency measure, such as conceptual and analytical thinking within the Problem-Solving category. Each sub-measure—in this case conceptual and analytical thinking—will need to be assigned a weighting within that measure—in this case Problem Solving. So, in this example, each could count for half, or 50%, of the Problem-Solving rating, which then counts for 20% of the overall Behavioral Competency rating. Behavioral Competencies account for 25% of overall teacher ratings.

All measures must be converted to or designed on the same numerical scale, so that the weighting percentages will calculate correctly, without bias toward scales with higher numbers.

### Special Issue: Considering Whom a Teacher-Leader Leads

The teachers a multi-classroom leader has on her team will affect student outcomes. A teacher-leader who leads three first-year teachers, or a team of persistently struggling teachers, will not likely achieve the same growth right away as a teacher-leader with a team of teachers who were already producing high-growth learning on their own.

Addressing the technical calculations are beyond the scope of this guide. However, the best evaluation systems for multi-classroom leaders will incorporate measures that consider how much higher the team teachers' student growth is compared with prior years (or compared with average growth for new teachers). You might consider this a "teacher's growth" in student growth. This will avoid discouraging teacher-leaders from leading teams of new teachers, which are especially common in high-poverty schools.

That said, Opportunity Culture schools are focused on achieving high growth with all students, every year. Team-teacher assignments and the specific roles teachers play on each team should be crafted to maximize the use of all team members' strengths and to minimize the impact of new and struggling teachers' weaknesses. The goal: high growth for all students served by every team. Assigning a multi-classroom leader to only struggling or new teachers can undermine the potential impact of the multi-classroom leader role and team structure.

Principals and district administrators can use the [Evaluation and Accountability Database Tool](#) and input into human resource information systems, both for communicating accountability accurately and for talent management. In nearly all cases, the data department within the district will need to make changes in how data are collected and reported to match Opportunity Culture roles. Ideally, the data department and human resource staff will also work with a vendor to provide a dashboard of Opportunity Culture staffing models, student outcomes, and other key data being reported by other Opportunity Culture schools.

Some data changes should be coordinated with the state, so that the district is not creating a new system for evaluation data collection and reporting while bearing the costs of maintaining the old system. The Opportunity Culture initiative leader and team can present options to the Opportunity Culture Evaluation Design Team or make decisions based on that team's prior input. The Opportunity Culture Evaluation Design Team should be kept informed of all decisions and be able to review and comment upon draft materials.

#### *Magnitude of impact, or "reach"*

To give teachers credit for extending their reach, evaluations should include a "reach measure" of the number of students for whom each teacher is formally accountable compared with a standard, one-teacher-one-classroom teaching role.

Such a reach measure may not affect the evaluation but rather the level on a career path of a given role, and the extra compensation attached to that role. Including the measure as part of an evaluation is really a communication tool. It reminds teachers that their effectiveness can be extended and celebrates the expanded impact that each teacher has in his or her advanced role.

Our suggested formula for districts sharing this measure: A teacher's impact = effectiveness X the number of students reached. Evaluations today focus only on the first part of the formula and not on the magnitude of impact, or "reach." Teachers who teach 15 students brilliantly are making a terrific contribution to those children. But one who teaches brilliantly and extends her reach directly or by leading peers with full accountability for her students' learning is contributing more and having a greater impact. Reporting reach as a multiple of the average one-teacher-one-classroom reach will show teachers this other, critical aspect of their contribution to students, the school, and their communities.

For example, an elementary teacher with a class of 20 students is teaching 100 percent of the average student load of 20, and would have a "reach score" of 1.0. But an elementary teacher with a class of 16 students is teaching about 80 percent of the average student load of 20, and might have a "reach score" of 0.8. One extending reach on a Time-Technology Swap team, without increasing

instructional group sizes, reaches approximately 133 percent of the average student load and would have a reach score of 1.33. Another, a multi-classroom leader leading a team serving four teachers' worth of students, or 400 percent of average, would have a reach score of 4.0. Showing that reach score alongside a teacher's effectiveness rating emphasizes for teachers that one way of improving and advancing their careers is to help more students successfully, which requires better planning, teamwork, leadership, and related skills.

Districts to date have used similar thinking developing career paths, and some have set student reach percentage increases as a requirement for some advanced roles. Requiring that elementary team teachers reach at least 33 percent more students than the average in traditional classrooms in the district is one example. Examples for multi-classroom leaders include: in the first level of leadership, reaching four times the typical number of students through team leadership; and in the second level, reaching at least six or seven times the typical number of students. States have an important role in encouraging reach, or magnitude of impact: Eventually, each state should track and report the percentage of students in each core subject and grade, overall and by student subgroup, who have excellent and effective teachers accountable for their learning.

States committed to reaching far more students with excellent teaching will eventually require LEAs to report **the percentage of students whose teacher of record is highly effective** by school and district, at least for each core subject, and **the percentage of students in various subgroups who have teachers at each level of effectiveness** (as determined by the approved teacher evaluation instrument). Note that in states where local districts each design their own evaluation system, it will be difficult to establish a sense of teacher effectiveness across the state. States could set goals for reach extension that increase over time, such as challenging districts to ensure that 75 percent of students experience instruction led by excellent teachers in all four core subjects, at a minimum, within five years.

In the model materials here, a place for indicating each teacher's reach is included at the top of the rating sheet for job responsibilities.

### *Matching formal accountability and data systems at district and state levels to roles.*

Principals must inform and collaborate with Human Resources and Opportunity Culture Initiative leaders in the **district** for matching teachers with the right students and team colleagues. After principals match teachers and teacher-leaders to the right students, subjects, and team teachers—the ones for whom each is responsible each year—districts must keep track of the decisions and communicate critical decisions to the **state**.

### Special Issues for Common Opportunity Culture Roles

Roles common in Opportunity Culture schools each have special content design issues:

**Multi-classroom leaders:** Multi-classroom leader evaluations must address the role of teacher-leader, not just teacher. They should include job responsibilities related to leading and managing the team, competencies for great leadership and team management, and outcomes of *all* the students served by the team.

MCLs should be “fully responsible” for all outcomes, even though teachers on the team may share responsibility for some or all of the students in one or more subjects:

- \* In practice, a multi-classroom leader should be responsible for all the student outcomes of all the teachers on his or her team.
- \* In formal systems, accountability should be 100 percent for the team leader and individual teachers, not a division of the full 100 percent between each teacher and the multi-classroom leader. This prevents distortion of the MCL's duty to the team when he or she teaches some students directly. The MCL should be just as accountable for all the students served by the team as each teacher is.
- \* Some teachers on a team may teach only certain subjects, and that should be reflected in the accountability plan. (See example, page 12.)
- \* Accountability must be formally recorded *by subject*, not just by grade.

- \* In no case should a multi-classroom leader be formally accountable for *certain* students served by the team *but not others*, or only for some subjects covered by the team—unless another team leader is formally accountable for the other subjects.
- \* Note that shared accountability gives everyone an incentive to do their best and to take their input into others’ evaluations seriously. Not being honest when a teammate, team member, or leader is struggling may harm one’s own outcomes and other colleagues’ outcomes. Being honest will lead to the development the person needs or to more appropriate placement in a role where the person can succeed.
- \* As discussed above, leading states and districts will incorporate prior growth of a teacher-leader’s teaching team into evaluations. This will prevent teacher-leaders from trying to avoid having new and struggling teachers on their teams. Best of all, this will be incorporated into staffing design choices about whom each teacher-leader has on a team—balancing teams with combinations of new/struggling and previously high-performing teachers—to help all teams produce high-growth student learning.

**Team Teachers’** roles will vary by:

- \* level—elementary or secondary;
- \* if and whether they extend reach; and
- \* the makeup of their team and amount of paraprofessional support.

The content of job responsibilities should reflect the subject expertise, teamwork, and student reach required of the role, the competencies to stay organized and connected when reaching more students and to collaborate with a team and/or paraprofessionals, and the outcomes for all students reached in each subject for which the teacher is actually responsible.

The expectations for effectiveness may vary by the role a direct-reach teacher holds:

- \* Teachers in advanced roles with substantial pay supplements based on a *highly effective* track record may have mentoring responsibilities and/or a higher bar of expected future student outcomes in order to remain at the same level. Their job expectations and competencies should

reflect any mentoring and other team responsibilities, such as advanced instructional planning, and outcome measures should align with the subjects and students covered and the district’s career-path requirements for advanced roles.

- \* Teachers in a role and at a pay supplement level designated for teachers with an *effective* teaching track record may need to maintain this track record to remain in a role extending their reach—alone or on a team—especially if a pay supplement is involved. Their job expectations and competencies should reflect any specific responsibilities, and outcome measures should align with the subjects and students covered and the district’s career path requirements.

Teachers on a team often will share responsibility for students, but sometimes the outcomes will be clearly divisible for specific students and subjects. When shared, responsibility may be 100 percent for multiple people or an equal percentage for each teacher adding up to 100 percent, depending on the formal accountability system.

In spirit, shared responsibility should be “all in”—100 percent—for every individual responsible for a student in a given subject, even though no one teacher is doing the job alone, as example on previous page shows.

**Paraprofessionals:** Paraprofessionals will typically serve a team, not just one teacher as many of today’s teaching assistants do. Their responsibilities also may include more independent supervision of student learning time—skills practice, projects, and digital instruction time, for example. Competencies will include a higher level of customer service orientation and flexibility than a typical assistant might need. And, depending on the role, they may have a large impact on students’ social, emotional, and behavioral development at school. All of these should be reflected in evaluation content, and ideally in a paid career path for paraprofessionals.

## Example: Team Teaching Accountability

Team teaching and teacher-led teams present special accountability design needs. In an accountability system, teachers should be matched to the students, subjects, and team members for whom they are responsible. Surveys of Opportunity Culture teachers show that correct matching is very important to teachers. When peer teammates share responsibility for the same students in a subject, ideally they will both be 100 percent accountable. They may split accountability evenly, 50 percent each, but this doesn't work well when team leaders or MCLs are part of the system. When responsibility is shared by a team member and an MCL or other team leader, it cannot be cut into percentages less than 100 percent without discouraging effective leadership.

- \* For example, four elementary teachers may cover a grade, with one of them serving as the team multi-classroom leader and teaching math part time. Among the team teachers, one may teach math and science to most of the students, and two may teach language arts and social studies (ELA/SS) to half the students each.
- \* The multi-classroom leader should be accountable for all students in all subjects, even if primarily teaching one subject to some students day to day.
- \* The math teacher should be accountable for all students he or she actually teaches, but only in math. The two ELA/SS teachers should be 100 percent accountable for a specific half of the grade's students each—the ones they actually teach—but only in ELA/SS. Or, if they both teach *all* the students—one teaching reading and the other teaching writing and social studies, for example—then they can both be 100 percent accountable for all the students in ELA/SS.
- \* Where an MCL shares accountability for certain students in a given subject with a team teacher—as the MCL and math/science teacher do—district and states should ideally call it 100 percent accountability for each teacher. So, the MCL may be 100 percent accountable for the grade's math/science outcomes, and the grade's math teacher also 100 percent accountable for the students (s)he teaches.
- \* Splitting accountability of multi-classroom leaders, or any team leaders who continue to teach directly, into percentages that add up to 100 percent is less effective. Most states were not considering this when designing accountability systems of the past, of course. Thus some states and districts currently divide accountability by the child. So, a student in math is 100 percent. If the student has two teachers accountable, then each teacher has 50 percent accountability. So, in the example above, the MCL would have 100 percent math/science accountability for the students he or she teaches directly, but only 50 percent accountability for the students served by the other math teacher on the team. This dilutes and distorts the MCL's responsibility as team leader, and discourages the leader's duty to develop and lead the team. The team leaders should select students to teach directly based on what will achieve the very best outcome for *all the students served by the team*. Over-incentivizing a focus on "my" students but not "their" students diminishes the value and power of team leadership.
- \* Here is the main concept: In performance evaluations, **both** teachers are accountable for the math teacher's student outcomes in math. However, the MCL is also accountable for the development of all teachers on the team and for student outcomes in all subjects. This accountability is not limited to students whom the MCL teaches directly.

## Tools

- Evaluation Input Sheet** from Step 2 offers a simple tool for getting stakeholder input into evaluation content and the process. The *District Evaluation Content Plan* provides a place to record final decisions.
- Districts should also review their emerging evaluation content plans against the *Lessons from Teacher Evaluation in Traditionally Staffed Schools*.
- Teacher and Staff Selection Toolkit* provides model job descriptions and competencies.
- Standard Evaluation Tools Package** provides simple tools for clarifying job responsibilities/reach, competencies, and outcome measures, and for evaluating teachers in each of these areas. They may be used for evaluating teachers and staff in Opportunity Culture schools and similar models (team-based, teacher-led, extended-reach roles). They provide space to include other measures

important in your school. Schools must adapt the materials to fit each school setting and role.

- Evaluation and Accountability Database Tool* is a checklist with guidance for building an accurate, useful database that matches teachers in schools using Opportunity Culture (or similar models) to the right roles, students, subjects, and teams.
- Annual Staffing Update Tool* is used to track staffing of Opportunity Culture positions and excellent/highly effective teachers, and calculate the percentage of students in each school who have an excellent teacher of record.

These materials are targeted for use with roles in six core Opportunity Culture school models. The materials may be used for positions in other, similar school models and roles. More detail is available on the school models included in this toolkit at <https://www.opportunityculture.org/a-complete-opportunity-culture/>.

---

<sup>1</sup> Spencer, L. M., & Spencer, S. M. (1993). *Competence at work: Models for superior performance*. Wiley.

<sup>2</sup> Hunter, J. E., Schmidt, F. L., Judiesch, M. K. (1990, February). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*. 75 (1), 28–42.

<sup>3</sup> Dan Goldhaber and Michael Hansen (2010). *Is It Just a Bad Class? Assessing the Stability of Measured Teacher Performance* CEDR Working Paper #2010-3 (Seattle: Center for Education Data and Research).