



Illinois State Board of Education

February 2013

Guidance Document 13-10

Guidance on Creating Operating Guidelines for Student Growth Models in Teacher Evaluation Systems

*document is intended to provide non-regulatory guidance on the subject matter listed above.
For specific questions, please contact the person(s) identified in the document.*

Dr. Christopher Koch, State Superintendent

Printed by AFL-CIO (AFSCME Local #288 and IFSOE Local #3236) Employees

Guidance on Creating Operating Guidelines for Student Growth Models in Teacher Evaluation Systems

Subject

Recommendations for implementing student growth models for teacher performance evaluations based on a set of basic guidelines for a growth model

Type of Guidance

This guidance document supports districts in understanding how to create operating guidelines for student growth models and providing a foundation for making decisions about the district's evaluation system.

Explanation

The Performance Evaluation Reform Act (PERA) states:

By no later than September 1, 2012, each school district shall establish a teacher evaluation plan.... The plan shall include a description of each teacher's duties and responsibilities and of the standards to which that teacher is expected to conform, and shall include at least the following components:... (c) by no later than the applicable implementation date, consideration of student growth as a significant factor in the rating of the teacher's performance. (PERA Sec. 24A-5)

The Illinois Administrative Code states:

The Joint Committee shall consider how certain student characteristics (e.g., special education placement, English language learners, low-income populations) shall be used for each measurement model chosen to ensure that they *best measure the impact that a teacher, school and school district have on students' academic achievement.*" (23 Ill. Adm. Code 50.110[c])

The Illinois State Board of Education (ISBE) [*Non-Regulatory Guidance*](#) states:

The student growth component of a teacher's evaluation may or may not "cover all students that the teacher instructs during his or her evaluation cycle.... School districts should strive to incorporate as many students that the teacher instructs as possible when incorporating data and indicators of student growth into a teacher's evaluation. (A-34, p. 26)

PERA, the Administrative Code, and the *Non-Regulatory Guidance* specify that:

- Student growth should contribute significantly to teacher evaluations.

- The characteristics of students that may affect the teacher’s impact on their growth should be taken into account.
- Teachers should be evaluated based on their impact on as many students as possible.

Assessments That Measure Student Growth

Notably, measuring student growth differs somewhat from traditional ways to measure student performance or attainment. Further, for a teacher to be evaluated fairly, the appropriate students must be identified so their growth will contribute to the teacher’s evaluation. In addition, student growth scores should contribute to teachers’ evaluation scores in a way that reflects the impact of each teacher on each student’s learning growth.

Consequently, a Joint Committee will need to do the following: (1) identify existing or new assessments that allow for the measurement of *student growth* in all subjects and grades where the state is not yet prescribing an assessment, (2) leverage existing data and data collection systems to correctly link students to teachers, and (3) design schemes that describe the extent to which each a teacher contributes to the learning growth of each student in each subject.

In addition, assessments are categorized into three types: Type I, Type II, and Type III (see Table 1). Each teacher will need to be covered by a Type I or Type II assessment as well as a Type III assessment. The state will provide guidance separately on the implementation of Type I, Type II, and Type III assessments.

Table 1. Illinois Assessment Types¹

Assessment Type	Description
Type I	An assessment that measures a certain group of students in the same manner with the same potential assessment items, is scored by a non-district entity, and is widely administered beyond Illinois (<i>Examples: Northwest Evaluation Association (NWEA) MAP tests, Scantron Performance Series</i>)
Type II	An assessment developed or adopted and approved by the school district and used on a district-wide basis that is given by all teachers in a given grade or subject area (<i>Examples: collaboratively developed common assessments, curriculum tests, assessments designed by textbook publishers</i>)
Type III	An assessment that is rigorous, that is aligned with the course curriculum, and that the evaluator and teacher determine measures student learning (<i>Examples: teacher-created assessments, assessments of student performance</i>)

This guidance document focuses on the decisions that are needed to be made *before* beginning to use Type I or Type II assessments to measure student growth. These decisions, or operating guidelines, are the basic questions that a Joint Committee should discuss with regard to measuring student growth and attributing it to teachers using Type I or Type II assessments. Though there are many more nuances to the following topics, the following guidance is intended to provide a foundation for the conversation.

¹ From the [Transforming Educator Evaluations in Illinois PPT](#).

Decisions, Recommended Guidance, and Examples

Decision 1: Determine which students are counted for a teacher’s growth measure.

A district can increase the reliability of a teacher’s evaluation scores by including the growth of as many students as possible in the score while ensuring that the teacher had a real opportunity to contribute to the growth of these students.

The district should include students who have both a pretest and a posttest score. Start from teachers’ rosters and recall that a pretest and a posttest are required to measure growth to determine whether the scores of some students cannot be used. The pretest can be from the end of the previous year during the previous grade, or from the beginning of the school year or course, depending on the assessment.

Any reason to exclude students other than missing test scores should be clearly documented and justified. In particular, it should be clear why the exclusion increases the fairness of the system for students and teachers.

Decision 2: Determine how long a student needs to be enrolled in or attending a teacher’s class for the teacher to contribute to his or her growth.

A student must have a pretest in order to contribute to his or her growth. The latest date a pretest may be given depends on the type of assessment; but, generally, the pretest must be given close to the beginning of the school year or course.

The time a teacher spends teaching a student may depend on the portion of the school year the teacher and student spend together. Often a student and teacher will be assigned to each other for a school year or a semester. Student absences also may affect the portion of the school year that the teacher and student actually spend together. These situations can be taken into account by assigning one of the three levels of interaction—primary, shared, or limited—based on the portion of the school days between the pretest and posttest that the teacher had responsibility for the present student(s). For example, a district will pick values for x percent and y percent, depending on their attendance patterns:

- More than y percent of days → primary
- Between x percent and y percent of days → shared
- Less than x percent of days → limited

Guiding Questions for Picking x and y :

1. What is the minimum portion of a school year or course that is necessary for a teacher to contribute to the learning of a student? What is the minimum portion of a school year or course that is necessary for a teacher to contribute *significantly* to the learning of a student?
2. What are the attendance patterns in the district? What is the average attendance? What is the minimum and the maximum attendance? How does this attendance differ across

elementary, middle, and high school levels? Does it vary in other significant ways? For example, a district where average attendance is 50 percent likely will pick different x and y than a district in which average attendance is 90 percent.

In order to account for these different categories, the district can assign weights to each link based on the length of time that a teacher teaches a student. For example:

- Primary level of interaction = Highest weight
- Shared level of interaction = Weight between 0 and 1
- Limited level of interaction = Lowest weight

Guiding Questions for Picking Weights:

1. What weights would accurately and fairly capture a wide range of teaching situations in the district?
2. Will growth calculations capture complex weighting schemes and affect growth results? (e.g., multiple weights for multiple subcategories of shared).
3. If a weight of zero is assigned to the limited responsibility category, what is the risk that teachers will pay less attention to some students?
4. Considering common coteaching situations in the district, what are those situations in which teachers equally share responsibility and those in which one teacher contributes much more than another? When two (or more) teachers contribute equally, when do they contribute equally *and* significantly?

For more information on this weighting scheme, see Decision 4 (“Determine how shared teacher responsibility for students, later teacher assignment, teacher absence, and/or teacher transfer are accounted for in student growth measures”) on page 5.

Decision 3: Determine the portion of a teacher’s students to be included in the growth measure.

A district can increase the reliability of a teacher’s evaluation scores by including the growth of as many students as possible in the score while ensuring that the teacher had a real opportunity to contribute to the growth of these students. In particular, this approach involves including all the students whom a teacher teaches; these students may be across different classrooms, grades, and content areas. When a teacher’s evaluation score is calculated based on the scores of too few students, the teacher’s score is less reliable. For example, the teacher’s score could drop significantly if one student performs poorly on a test because he or she is not feeling well on test day and there are only eight students in the classroom.

Because all teachers must have student growth included in their teacher evaluation, regardless of the number of students they have, the district and union can consider using more than one year’s worth of growth data for a teacher when that teacher has a low number of students.

Guiding Questions for Picking a Minimum Number of Students:

1. How many students does it take for an average score to reflect a teacher’s contribution to their learning?
2. How many students does it take so that the average score is not overly changed by an unusually low or high performance by one or a very small number of students?
3. Is there a certain threshold (such as a minimum of 10 students), under which the student growth of a teacher would be calculated using more than one year of data?

In Illinois, not calculating or not reporting a student growth-based score for a teacher with too few students, and increasing the weight of other measures into their overall evaluation score instead, is not an option. In addition, all students whose data are included in a teacher’s score must be assigned to that teacher. The district and union can consider assigning a score based on a larger group of students to those teachers. For example, starting in the second year of implementation, scores could be based on the growth of students taught during the course of more than one year, thus increasing the sample size and the reliability of results.

Decision 4: Determine how shared teacher responsibility for students, later teacher assignment, teacher absence, and/or teacher transfer are accounted for in student growth measures.

Different teachers may spend more or less time with different students in their classroom(s), depending on scheduling and teaching arrangements. Likewise, the time that a teacher spends teaching a student depends not only on the teaching or coteaching situation but on the portion of the school year that the teacher and student spend together. Often, a student and teacher will be assigned to each other for a school year or a semester; yet the teacher may be assigned to a different classroom later in the semester or school year or the teacher can be absent sometimes for significant lengths of time.

A district and union can ensure that teachers’ evaluation scores reflect the different contributions of different teachers to different students’ learning growth by assigning one of the three levels of responsibility—primary, shared, or limited—based on the level of interaction between the student and educator in the subject area being measured or the portion of the school days between the pretest and posttest in which the teacher had responsibility for the students. These mirror the levels of interaction described in Decision 2 (“Determine how long a student needs to be enrolled in or attending a teacher’s class for the teacher to contribute to his or her growth”).

Interaction Between Student and Educator:

- **Primary responsibility** indicates that the teacher is principally responsible for that student’s learning in the content or standards being assessed.
- **Shared responsibility** indicates that the teacher is partially but not fully responsible for that student’s learning in the content or standards being assessed (i.e., that responsibility is distributed among two or more educators).
- **Limited responsibility** indicates that the teacher has insufficient responsibility for the student’s learning for that student’s assessment results to be used in any significant way in determining the teacher’s evaluation score.

There probably are too many types of coteaching situations in the district to list them all. Curriculum and Instruction staff will create a few categories that summarize existing schemes into a small number of categories into which each teacher feels they can fit. For example:

- **Primary Responsibility:**
 - I was the only teacher for this student in this class.
 - I provided most of the instruction to this student in this class, but the student did receive additional support from another teacher.
 - Another teacher and I had joint responsibility for the learning of this student in this class; one of us is a general education teacher, and the other is a teacher of special education.
- **Shared Responsibility:**
 - Another teacher and I contributed significantly to the learning of this student in this class.
- **Limited Responsibility:**
 - I provided support to this student in this class.

Then, portions of school days and weights can be assigned as described above. Each district can fill the following matrix (see Table 2) with its choices of portions of school days and weights.

Table 2. Matrix on Categories and Weights for Teacher-Student Links

Teaching or Coteaching Situation	Time That Student Spends With Teacher		
	More than y% of Days	x%-y% of Days	Less than x% of Days
I was the only teacher for this student in this class.	Primary (Highest weight)	Shared (In-between weight)	Limited (Lowest weight)
I provided most of the instruction to this student in this class, but the student did receive additional support from another teacher.	Primary (Highest weight)	Shared (In-between weight)	Limited (Lowest weight)
Another teacher and I had joint responsibility for the learning of this student in this class; one of us is a general education teacher, and the other is a teacher of special education.	Primary (Highest weight)	Shared (In-between weight)	Limited (Lowest weight)
Another teacher and I contributed significantly to the learning of this student in this class.	Shared (In-between weight)	Limited (Lowest weight)	Limited (Lowest weight)
I provided support to this student in this class.	Limited (Lowest weight)	Limited (Lowest weight)	None (Weight = 0)

Decision 5: Ensure that student rosters are accurate and the correct students are being included in a teacher’s growth measure.

Before a teacher’s score can be constructed from several student scores, students must be assigned to one or more teachers so student growth scores can contribute to teachers’ evaluation scores, and decisions must be made on what factors will affect the amount of contribution of teachers to student learning. Decision 3 and Decision 4 assumed that information on teacher-student links is available and of high quality. If this information is not available and of high quality or if it is not clear that it is, a process of roster verification and information collection may be warranted.

The state is not requiring that districts put in place a formal system of roster verification, but districts are encouraged to ensure the accuracy of the information in their systems so evaluation scores are as accurate and fair as possible.

First, to verify the accuracy of existing information, the district can cross-examine data from different sources to look for discrepancies that may indicate data-entry errors.

- Audit a small number of random schools; create a simple Excel file requesting information on teacher-student links and compare the data to those found in existing administrative data. If the discrepancies are few, existing data systems are reliable.
- If the discrepancies are great, convene data and accountability staff to provide solutions, such as streamlining data collection processing or training relevant school staff.
- Communicate to schools the consequences of mistakes in data collection. If school staff understand how important these data are and how the data can affect their teachers, they will have an incentive to report the data as best they can.

Second, the district can identify what data are needed to correctly attribute students to teachers and what variables are not currently being collected. Then the district can put in place processes to fill the gap. Options include the following:

Create Teacher-Student Links Using Course Information

More than one class/course may use a specific assessment as its measure of student growth. As a result, the scores of students on one assessment may count toward the evaluation scores of more than one teacher across different courses. For example, a journalism course will contribute to a student’s performance on the Language Arts section of either of the Educational Planning and Assessment System assessments (Grades 8–9 EXPLORE assessment or Grade 10 PLAN assessment). It is therefore appropriate to use the EPAS growth scores as the Type 1 assessment for a journalism teacher. That teacher also will have a Type 3 assessment that is directly tied to the course content.

The district will connect student achievement on tests to teachers of corresponding courses. Therefore, it will identify which courses students took that prepared them for these tests and which teachers taught them during those courses. First, the district will list available assessments. Then, for each assessment, curriculum specialists will identify which courses prepare students for it. Next, data teams will list teachers for each of these courses. Finally, they will list students for each of these teachers. Each student who took the pretest and posttest will count toward that teacher’s evaluation score.

Note that part of this process will include understanding the organization of homerooms, which may differ from one school to the next. A teacher may teach reading/ELA and mathematics in a homeroom—in which case, student scores on both those tests will count toward the teacher’s evaluation. Two homeroom teachers may decide to specialize, with one teacher teaching reading/ELA to students in both homerooms while the other teaches them mathematics. In this case, the first teacher will be evaluated based on the reading/ELA assessment, while the other will be evaluated based on the mathematics assessment.

- Ask principals to identify a person on staff in the school who can accurately report on teacher-student links, teacher assignments, and coteaching situations. Design a form that these individuals can fill out to provide the information needed for attribution. Require that all attribution lists are checked for accuracy by the teacher(s) to which the students have been attributed.
- Provide teachers with Excel lists of the students assigned to them in administrative data, and ask them to verify those lists; ask principals to sign off. Include columns that allow teachers to provide coteaching information. Ask principals to gather teachers during a staff meeting to fill the forms together and make sure they are in agreement.
- Implement a formal roster verification process. This process can be created in-house in collaboration with curriculum and instruction, data and accountability, human capital, data systems, and other relevant teams. Leverage publicly available information and tools that are available free or for purchase.²
- In all cases, provide teachers with an opportunity to double-check final rosters before they are uploaded and used to calculate scores.

Decision 6: Determine a course of action for when assessment data are missing for a student or group of students.

A student who does not have a score for either the pretest or the posttest will not be counted in the teacher’s growth measure. Reasonable efforts should be made to obtain missing assessment information. For example, if a student transferred into the school but took the assessment used as a pretest in his or her prior school, the score from the sending school may be used as a pretest. If it is early in the term, the student also may be allowed to take the pretest. If the pretest and/or posttest scores are missing for a student and cannot be obtained, that student is not included in the teacher’s growth measure.

Decision 7: Determine if students who skipped a grade or are held back a grade should be excluded from the growth attributed to a teacher.

Students who either skipped a grade or were held back should be included in the teacher’s student growth measure if they took the pretest and posttest used to measure growth. Regardless of how a student arrives in a teacher’s classroom, he or she should be expected to grow. The challenge with students who either skipped a grade or are held back is to ensure that they have taken the appropriate pretest if it was administered at the end of the last academic year.

Decision 8: Determine how to address the link between subjects and courses and particular assessments.

A Joint Committee should begin by identifying high-quality assessments with the appropriate properties for measuring growth. Some of these assessments may be prescribed by the state; others will need to be purchased or created, until the state has prescribed assessments for all

² See, for example, [Battelle for Kids guidance and tool for roster verification](#), or resources related to the Center for Educational Leadership and Technology’s pilot of [Teacher-Student Data Link](#), a roster verification process, in five states.

grades and subjects. These assessments need to provide achievement data over time; each student should receive at least a pretest and posttest score, with the pretest for the previous year and grade or from the beginning of the school year or course for that grade and subject.

- First, involve district staff with knowledge of curriculum and assessment to establish a list of the properties of high-quality assessments, including but not limited to reliability, validity, and alignment with appropriate standards and course content. Conduct a survey of assessments already in use in the district. Establish a list of known assessments in the nation. Identify grades and subjects where new assessments will need to be created because no assessments exist or because existing assessments do not have the appropriate properties.
- Next, examine the properties of assessments both for quality and for the ability to measure growth. Information on the properties of off-the-shelf assessments is generally available from vendors. Develop new assessments, keeping in mind the properties of a good assessment; ensure that the assessments are to be used for teacher evaluations, that they must measure growth over time, that they must be comparable across schools, that they must be appropriate for all students to be tested regardless of cognitive ability, and that they can be administered securely. If possible, pilot new assessments to verify their properties and modify them if needed.

Frequently Asked Questions

1. Does our district have to follow what is outlined in this guidance document?

No. This is a guidance document. In the development of the teacher evaluation system, a Joint Committee, in collaboration, can choose whether to follow this guidance. Teachers who have Type I and Type II assessments available must use them in calculating their growth score, however.

2. What is the timeline for implementation?

All school districts statewide must implement new teacher-performance evaluation systems, based in part on student growth measures by September 1, 2016—except for the lowest performing 20 percent of school districts across the state which must implement by September 1, 2015, and Chicago Public Schools, which must implement by September 1, 2013. Please visit <http://www.isbe.state.il.us/peac/html/timeline.htm> for more details.

3. What role will the state play, if any, in reviewing and/or vetting the assessments that our district purchases or creates? What information will we need to provide the state to document the quality of chosen assessments?

The state has been developing guidance on creating and enacting new educator evaluation systems and how to gauge their quality. More information is forthcoming.

4. Is a district solely responsible for calculating student growth scores for its teachers and administrators?

Yes, each district is responsible for calculating student growth for its teachers and administrators. The state does not provide direct support in growth calculations or data management. There is no statewide longitudinal growth model.

5. Does a district have to submit its final evaluation plan to the state for approval?

Not at this time.

6. How do administrators and teachers share their views?

Teachers and administrators will have multiple chances to share their views. First, the public is welcome to attend monthly meetings of the Performance Evaluation Advisory Council (PEAC) as recommendations are crafted. Meeting dates and more information can be found at <http://www.isbe.net/PEAC/>.

Second, PEAC is seeking input from teachers, administrators, parents, and the public through a series of meetings across the state, as well as through online surveys.

Please visit <http://www.isbe.net/PEAC/> to learn more about how you can help transform education in Illinois.