

**Comment and Review Form
Proposed Regulations on Evaluation of Educators**

(Submitted by AFT Massachusetts, June 10, 2011)

Name and title of person completing this form:

Thomas J. Gosnell on behalf of AFT Massachusetts

Position and/or discipline:

President, AFT Massachusetts

School district or affiliation: N/A

Email:

tgosnell@aftma.net

This review was completed by: an individual a group (please specify):

These comments reflect the collective work of leaders, members, and staff at AFT Massachusetts.

Your Primary Role:

PreK-12 Administrator (please specify):

Teacher (please circle grades) PreK K 1 2 3 4 5 6 7 8 9 10 11 12

Higher Education (please specify):

Other PreK-12 educator (please specify):

Other (please specify): State teacher union leader

Please feel free to add or duplicate pages so that you have enough room for your comments.

Overview

Dear Commissioner Chester and members of the Board of Elementary and Secondary Education,

Thank you for the opportunity to comment on the proposed educator evaluation regulations, 603 CMR 35.00.

In addition to submitting this form, **we are also submitting separately a marked-up version of the draft regulations. The marked-up version contains our full set of edits.** Some of the edits are copy edits or are minor edits designed to improve clarity and coherence; these edits are for the most part self-explanatory. Other edits are more substantive in nature and are therefore explained fully in the comments below.

At AFT Massachusetts, we believe that all education reforms should be good for students, fair to educators, and collaboratively done. That core belief runs through our comments and edits.

In terms of educator evaluation specifically, we support a system that emphasizes continuous improvement in a climate of trust and respect—not “gotcha.” We believe that our edits help us get closer to that goal, creating a system that will benefit students, parents, and educators alike.

We are happy to meet with you at any time to discuss our comments and edits in greater depth. Thank you again for your time and consideration.

Please address comments within the following categories:

I. Section 35.01 - Scope, Purpose, and Authority

Comments:

II. Section 35.02 - Definitions

Comments:

Developing Teacher/Administrator Plan and “progress toward goals”

The proposed regulations require educators to set professional practice goals and student learning goals. “Progress toward goals” can then be a factor in the summative rating, per the recommendation in the state’s Educator Evaluation Task Force report (p. 22), and per the definition of “summative evaluation” in the proposed regulations. However, the phrase “progress toward goals” is not used consistently throughout the proposed regulations. We therefore inserted this edit throughout the document, starting with the definition of “developing teacher/administrator plan.” The same change was made for the definitions of directed growth plan, improvement plan, and self-directed growth plan.

As the Task Force discussed and unanimously agreed to, the phrase “progress toward goals” is superior to language such as “must meet goals” for the following reasons:

- It enhances, rather than restricts, the professional judgment of evaluators.
- It ensures that educators who set high goals but fall short of them need not be automatically penalized but can instead get “credit” for progress. This creates a safe space for educators to set ambitious, challenging goals instead of low goals that are easy to meet.
- It acknowledges, especially in the case of student learning goals, that external factors may sometimes impede the attainment of goals, despite the best efforts of educators, students, parents, and others.

Directed Growth Plan

We are recommending that directed growth plans be for one year. Therefore, we struck the language “or less” from the proposed regulations. We make this edit for the following reasons:

- This will ensure clarity and consistency across districts in how directed growth plans are implemented.
- The primary purpose of a directed growth plan is to help educators get better so that they can reach proficiency. For example, an experienced teacher may not necessarily be “proficient” right off the bat if his/her assignment changes and the teacher is suddenly facing a new grade, subject, curriculum, or student population. This doesn’t mean this person is a “bad” teacher; it just means he or she needs some time and assistance to get up to speed. Establishing directed growth plans for a full year will help assure educators that these plans are about improvement, not “gotcha,” and will reduce disincentives for educators to take on new assignments and challenges. (For these same reasons, we oppose the proposed language in 35.08(8) saying that “an educator who receives a needs improvement rating shall receive this rating for one year only.” This kind of formulaic, automated approach restricts professional judgment and creates disincentives for educators to consider taking on new challenges and assignments. We’ll address this problematic language in greater detail in the appropriate section.)
- If administrators believe that an educator warrants a plan of a shorter duration, they have the option of rating the educator “unsatisfactory” and placing the educator on an improvement plan.

Evaluation—summative evaluations vs. formative assessments

To clearly distinguish between these two concepts, we recommend deleting the word “evaluation” after “formative assessment” wherever it appears. Formative assessments should center on feedback, support, improvement, and growth; summative evaluations result in a rating that carries consequences. We made this edit throughout the document, starting with the definition of “Evaluation.”

Goal

Throughout the proposed regulations, we added the following language: “Goals may be developed individually, or by school-level, department-level, grade-level, or other teams of educators who share responsibility for student results.” This language acknowledges what research has proven: Collaboration among educators is essential to school and student success. We must develop policies and systems that explicitly promote collaboration among teachers, administrators, and other school personnel. The inclusion of this language throughout the regulations will send the strong message that districts and schools are not only allowed to create group goals but are encouraged to do so.

Performance Standards

Throughout the proposed regulations, we delete references to performance “indicators.” We do not believe that the Department has the authority to mandate indicators at the local level. We also do not believe that it is a good idea to do so. Districts should have the flexibility to use long-established and successful evaluation rubrics, such as Kim Marshall’s or Charlotte Danielson’s, and **the regulations should not preclude the use of such rubrics, especially in districts that have invested significant resources and training into their use.** Having said that, it is appropriate for the Department to include indicators and descriptors in its voluntary “model system.” If this model system is designed thoughtfully and with educator input, we suspect that many districts will be inclined to use it.

III. Section 35.03 - Standards and Indicators of Effective Teaching

Comments:

Per our earlier comment, our major concern here is that districts should retain the flexibility to use long-established and successful evaluation rubrics. Therefore, we revised the language to say that school committees shall establish evaluation systems that “are consistent with the four core standards adopted by the Board.”

As mentioned earlier, we deleted the word “indicator” throughout. We also deleted all the descriptive information under each indicator, as this language is overly prescriptive. Such descriptive information would be appropriate, however, in the state’s voluntary “model system,” which local districts could choose to adopt or adapt.

We also revised the family and community engagement standard to eliminate roles and responsibilities that go beyond those of the typical teacher. A teacher’s work should be centered on engaging families, caregivers, and school-related organizations. Engagement with the broader community may occur as part of a teacher’s job, but it should not be a central component of an evaluation rubric.

IV. Section 35.04 – Standards and Indicators of Effective Administrative Leadership

Comments:

Consistent with our comments on 35.03, we revised language to allow more flexibility in the use of established, successful evaluation rubrics.

V. Section 35.05 – Evaluation of Administrators under Individual Employment Contracts

Comments:

Per our earlier comments, we removed language relating to indicators.

VI. Section 35.06 – Evaluation Cycle

Comments:

(1) School committee requirements

We deleted language relating to a “comprehensive evaluation system designed and regularly updated by the Department,” as that option is addressed in 35.10 (3).

(2) Self-reflection and self-assessment

It is confusing and probably unnecessary to refer to both words here. Since self-reflection is a natural part of self-assessment, it seems that “self-assessment” alone would suffice here.

(3) (a) Educator Plans—Goals

Per our earlier comment, we added language to make it clear that goals may be set individually or collectively.

(3) (b) Education Plans—Support

We added language saying that plans shall be designed to provide educators with “adequate time, support, resources, and feedback.” This reflects our core belief that accountability must be a two-way street. Just as educators will be held accountable for completing the actions in their plans, districts and schools must be held accountable for providing the time, resources, and support that will help educators be successful.

(3) (c) Educator Plans—Types of plans

We edited this section to conform with the near-unanimous recommendations of the Task Force report:

- All non-PTS educators rated proficient or higher be placed on Developing Plans
- All experienced educators rated exemplary or proficient be placed on Self-Directed Growth Plans
- All experienced educators rated needs improvement be placed on Directed Growth Plans
- All experienced educators rated unsatisfactory be placed on Improvement Plans

We added language requiring educators with “low” impact on student learning to develop growth plans that include strategies designed to increase student learning.

(3) (d) Educator Plans—Requirements

We deleted language saying that student learning goals must “be measured by improvements in either statewide assessments or district-wide assessments.” In its place, we added language

saying the goals would be “assessed through one or more of the student performance measures set forth in 35.07(1)(a).” Please note that we also rewrote 35.07(1)(a) to include “progress on teacher-developed or teacher-selected assessments that are approved by the district or school and that use a common scoring/grading rubric across similar classrooms.”

Our major concern here is that the proposed DESE language, which limits measures to either “statewide assessments or district-wide assessments” is far too restrictive and likely unworkable. Consistent with the views of Task Force members, we urge a broader conception of student learning measures here and throughout the document. Consider the following points:

- The primary statewide measure is the MCAS growth measure (median SGP). Aside from the fact that this measure applies to less than 20 percent of teachers statewide, we continue to believe that this measure is an inappropriate source of evidence for assessing teacher performance. The MCAS was not designed or validated for this purpose, and using the test in such a manner could create numerous unintended consequences, such as narrowing the curriculum, assessing teacher performance inaccurately, and discouraging great teachers from serving the neediest schools and students.
- The “district-wide assessments” referred to in this section of the regulations do not yet exist and may never exist for many groups of educators (such as special education teachers, art teachers, physical education teachers, and guidance counselors). Moreover, developing and administering tests in every grade and subject for educator accountability purposes will carry huge costs for districts and will likely come at the expense of other district priorities.
- Given these limitations and drawbacks, the better approach is to give front-line educators significant input into developing the student performance measures by which they will be held accountable. Our rewrite of 35.07(1)(a) and the inclusion of the rewritten 35.07(1)(a) in this section creates that possibility while honoring the Race to the Top requirement that measures be “rigorous and comparable” across classrooms.
- Our changes are consistent with the Task Force recommendations. The Task Force deliberately elevated “measures of student progress and/or achievement toward student learning targets set between educator and evaluator for the academic year” as the number-one category of evidence under “multiple measures of student learning, achievement, and growth” (p. 17). The Task Force clearly saw a large role for educator-developed or educator-selected measures in a true “multiple measures” system. Yet, under the proposed regulations, educator-developed measures play little or no role. Our re-write restores the role of educator voice and expertise in the development of multiple measures.

Per our earlier comments, we added language making it clear that goals may be set individually or collectively.

(4) Implementation

Per our earlier comment that accountability must be a two-way street, we added language saying that “the school district is responsible for ensuring that the educator has the necessary time, resources, and supports to complete the actions in the plan,” and that “all prescribed PD activities

shall be at no cost to the educator.”

(5) Formative assessment

We deleted the word “evaluation” after “formative assessment,” per our earlier comment.

As written, the proposed regulations do not make a clear enough distinction between the formative and summative stages of the cycle and the different purposes of each. We therefore revised the paragraph to make it clear that formative assessment should center on improvement-oriented feedback. Consistent with that focus, we do not believe that evaluators should give an “initial rating” during or upon completion of the formative assessment, and we deleted such references.

(6) Summative assessment

We added the phrase “progress toward goals,” per our earlier comment.

We also added language saying that “in assessing progress toward student learning goals, the evaluator shall consider school/classroom conditions, the availability and degree of classroom resources and supports, and other factors beyond the control of the educator that may influence student learning.”

We believe this language is essential to creating a process that educators perceive as credible and fair. The change is also consistent with recommendations from the Task Force. Here are two excerpts on this matter from the Task Force report:

“Many on the Task Force believe that disparities in conditions or resources must be taken into account during the evaluation process.” (p. 12)

“Conditions and resources needed by the educator to meet state Standards should be considered by the evaluator, where appropriate.” (p. 18)

(7) Placement of educators on plans

We revised this section to be consistent with the Task Force recommendations and to be consistent with our rewrite to 35.06 (3)(c).

VII. Section 35.07 – Evidence Used in Evaluation

Comments:

Per our earlier comments, we revised (1)(a) to include “school- or classroom-based learning goals” that may include “progress on teacher-developed or teacher-selected assessments that are approved by the district or school and that use a common scoring/grading rubric across similar classrooms.”

For the record, and for the reasons stated earlier, we remain opposed to the use of MCAS as a measure of teacher effectiveness. However, if MCAS must be used, it is imperative that evaluators look at **trends over time**, not a one-time snapshot.

As stated earlier, we remain unclear on what “district-determined measures” will look like and how they will be developed and used. In the absence of such information, we urge the inclusion of educator-developed measures (as written above) in a true multiple measures system.

VIII. Section 35.08 – Performance Level Ratings

Comments:

(2) Rubrics

We believe that our rewrite strikes a better balance between uniformity and flexibility on the issue of how districts will design or select rubrics that adhere to the regulations.

(3) Determining Ratings

Per our earlier comment, we added language saying “in assessing the educator’s impact on student learning, the evaluator shall consider school/classroom conditions, the availability and degree of classroom resources and supports, and other factors beyond the control of the educator that may influence student learning.”

(8) “Needs Improvement” for One Year Only

We deleted this provision because we strongly disagree with it. Imposing such a formulaic decision-making process on districts restricts professional judgment and will likely discourage experienced educators from taking on new assignments and challenges. Make no mistake: We recognize that subpar teaching cannot go on indefinitely and that there will certainly be cases where a teacher rated “needs improvement” would warrant being moved to “unsatisfactory” at the end of the directed growth plan. But we can also foresee cases—for example, that of an experienced teacher who has “stepped up” to teach a high-need subject, field, grade level, or student population—in which an additional year on a directed growth plan may be the appropriate path to help an educator reach proficiency. Our major argument is that these decisions should be left to the judgment of the evaluator and school district, not dictated through one-size-fits-all regulations.

IX. Section 35.09 – Student Performance Ratings

Comments:

(2) Impact on Student Learning Rating

This is the most confusing and problematic section of the proposed regulations. It's important to note that the concepts contained in this section of the regulations were considered and overwhelmingly rejected by the Task Force at the February 10, 2011 Task Force meeting. The notes from that meeting (copied and pasted verbatim below) are clear on this point:

“The following proposals were all clustered as a part of the Gates and Threshold discussion:

Proposal: Districts must differentiate educator contribution to student learning & growth

Proposal: Districts must develop a plan to differentiate educator contribution to student learning & growth into three levels

Proposal: Districts must include multiple measures in their Student Learning & Growth rating

Proposal: Specify required “thresholds” or “gates” for learning and growth in the framework

Strong feeling expressed: Feels like a last effort to use MCAS as more than one measure. It feels as if this was created by department, not by Task Force. Feels like something else is moving this forward, rather than will of Task Force. If this idea gets brought up again next week, it will reflect on that the consensus of the Task Force was not accepted. The end of this memo is toxic.

By general agreement, none of the gating/threshold proposals had support.”

Moreover, on April 22, 2011, AFT MA President Tom Gosnell sent a memo to Commissioner Chester asking for clarity on this and other sections of the proposed regulations. To date, we have not received answers to fundamental questions, such as:

- why a separate “impact on student learning” rating is needed at all when student performance measures are already factored *twice* into the “summative” rating: 1) as evidence to assess performance against the standards; and 2) as evidence to assess progress toward growth plan goals; what’s the rationale for “counting” student measures a third time?
- how districts will arrive at this “impact on student learning” rating in a consistent and reliable way;
- how “high,” “moderate,” and “low” will be defined when tools for measuring student growth against a criterion-referenced standard (such as “one year’s growth”) do not exist (even the MCAS growth measure can only provide a **norm-referenced** distribution of student/educator scores);

- how attribution at the individual teacher level will be done accurately when there are often multiple educators or multiple factors contributing to student learning simultaneously;
- how student assignment patterns will be factored into judgments; for example, what will be done to ensure that teachers assigned a disproportionate share of students with academic or test-taking challenges are not unfairly penalized under this system?
- what “district-wide measures” might look like for different educator roles, particularly those that don’t lend themselves to conventional assessment (e.g., ELL teachers, special education teachers, music teachers, physical education teachers, etc.).

For all these reasons, the most prudent course of action would be to put this section “on hold” — at least until some of these fundamental questions are answered. Short of that option, we urge a substantial rewrite of 35.09 to address these concerns, and we provide language in our edits. We believe that our rewrite does a better job of constructing a true multiple measures system that honors educator voice and ensures fairness.

(3) Resolving rating discrepancies

We rewrote parts of this paragraph in the interest of creating a fairer and more transparent process.

X. Section 35.10 – Implementation and Reporting

Comments:

We deleted 35.10 (4), as our rewrite of 35.09(2) makes it obsolete.

XI. Other General Comments:

As noted at the outset of our comments, an educator development and evaluation system should be about continuous improvement, not “gotcha.” We strongly support the concept of continuous improvement in a system that is fair, transparent, respectful of educator voice, and built on trust and collaboration.

But just as this concept applies to educators, it must also apply to public policy. In many ways, these regulations are venturing into unknown territory, and we need to maintain a sense of humility about what we know and don’t know about how to design a fair and educationally sound system.

As we move forward with implementation, it is essential that all stakeholders commit to careful monitoring and continuous improvement of the system as new evidence and research arises. This includes research from high-performing nations, many of which take an entirely different approach to educator development and evaluation, as well as to high-stakes testing and school accountability.

For a new system to work well, stakeholders must also seek continuous input from the educators on the front lines who will be directly affected by the changes. Educators’ voices must be heard and heeded throughout the ongoing process of implementation, monitoring, and refinement of the system.

We are hopeful that we can design and sustain a 21st century educator development and evaluation system that is both good for students and fair to educators, but this is possible only if we abandon pre-conceived agendas, follow the evidence and research, and work collaboratively in pursuit of common goals. We embrace this challenge and the opportunity to build a system that truly empowers front-line educators to take charge of their profession.

