MET project

POLICY AND PRACTICE SUMMARY

What It Looks Like

Master Coding Videos for Observer Training and Assessment

Catherine McClellan, Ph.D., Clowder Consulting, LLC

Assessment of classroom teaching that results in meaningful and actionable feedback for teachers has its roots in a clear, well-designed observation rubric. A number of research-based rubrics have come into use to frame such observations around specific aspects of practice believed to have the strongest positive impact on student learning. These instruments break teaching into discrete components (e.g., "Use of Questioning," "Behavior Management," and "Modeling") and, for each component,* describe the observable attributes of different levels of performance (e.g., "unsatisfactory," "developing," "effective," and "highly effective").

But without accurate, reliable application of these tools, the potential improvements in teaching and student learning will be lost, buried under a jumble of rater bias, observer errors, and extraneous factors that have little or nothing to do with a teacher's demonstrated proficiency. The resulting lack of consistency in scores will erode trust in the whole enterprise of teacher evaluation.

Observers must know what it looks like when a particular aspect of teaching is demonstrated at a particular level. An increasingly important tool for ensuring this competency is master-coded videos—videos of teachers engaged in classroom instruction that have been assigned correct scores by people with expertise in both the rubric and teaching practice. These videos become the benchmarks for the development and assessment of observers (see **Figure 1**). They also can be used to help teachers better understand their system's expectations for effective teaching. Without such benchmarks, an observation rubric remains a well-organized set of written definitions and "look-fors" that could (and likely would) be interpreted by different individuals in different ways.

Rubric developers, consultants, and vendors often can provide master-coded videos to school systems adopting a particular instrument. But when a locally developed instrument is used, when an existing one has been modified, or when examples from the local context are especially important, states, districts, or consortia of districts may need to take on master coding themselves. Even those who rely primarily on videos master coded by others can benefit from becoming educated consumers of them.

Beyond the value of producing coded videos, the engagement of principals, teachers, and peer observers in the master coding process can help foster a shared agreement about what effective teaching looks like, which is essential to the buy-in and success of any feedback and evaluation system.

Beyond the value of producing coded videos, the engagement of principals, teachers, and peer observers in a master coding process can help foster a shared agreement about what effective teaching looks like, which is essential to the buy-in and success of any feedback and evaluation system. Master coding can help clarify a rubric's language, and it can help those who code provide better feedback when they work with teachers.

Benefits aside, master coding is challenging. It requires careful planning, recruitment, training, and management. Anyone involved in organizing a master coding effort must understand the essentials outlined in these pages.

©2013 Bill & Melinda Gates Foundation.
All Rights Reserved. Bill & Melinda Gates Foundation is a registered trademark in the United States and other countries.

BILL & MELINDA GATES foundation

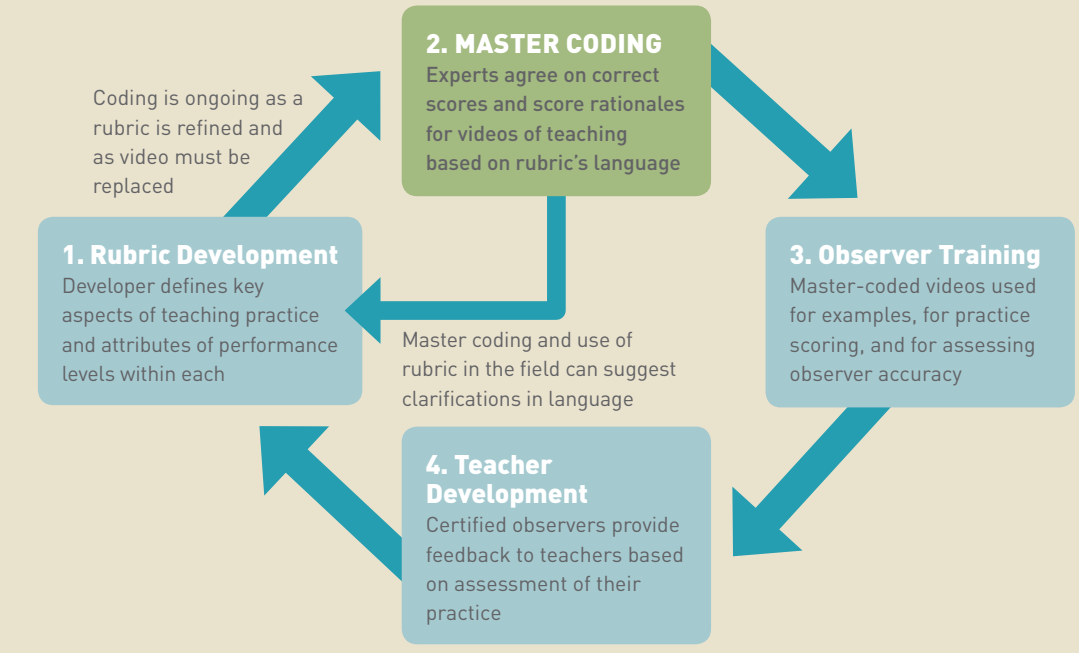
THE GOALS OF MASTER CODING

Master coding supports consistent application of a rubric. The process involves individual scoring of raw videos of classroom instruction, followed by discussion, often in pairs, to reach consensus. To give concrete meaning to an instrument's written descriptions, this process should produce the following.

- **Coded video that covers all of a rubric's components.** These videos include:
  - A series of short segments (e.g., two to seven minutes), each of which exemplifies one teaching component at one performance level. In observer training, these are used to illustrate a rubric's components and to call out their critical attributes.
  - Extended segments (e.g., 20 to 30 minutes) that include examples of multiple components and performance levels. Observers-in-training score these examples to practice and to demonstrate their accuracy.
- **Codes that make specific connections between the video and the rubric.** They must include:
  - Time stamps indicating at which points in a segment a particular component is illustrated.
  - Scores for the component exemplified.
  - Rationales, written in the language of the rubric and citing objective evidence, for why the observed behavior merited a particular score.
  - Rationales for why the behavior did not merit higher or lower scores, again using the rubric's words.

Figure 1

MASTER CODING IN A CLASSROOM OBSERVATION SYSTEM



* Although "component" is used here, various terms are used in the field—often inconsistently—to refer to the level of organization of a rubric at which scores are assigned. Other words include "standard," "dimension," and "element." Measurement experts typically use the word "scale."

- **Feedback for rubric refinement.** Master coding can bring to the surface the need to clarify an instrument's language and distinctions. When this happens, recommendations for refinement should be made to those with the authority to make such changes.

RECRUITMENT OF MASTER CODERS

Master coders are typically principals, teachers, and peer observers who may devote perhaps 12 hours a month to the process after their initial training. Depending on a rubric's complexity and the supply of quality raw video available, it could take two dozen part-time coders one year to code a full complement of video segments for a rubric. After that, coding should continue as videos become outdated and new segments are needed for observer assessment.

Given the commitment required, master coders should be selected for qualities most likely to ensure success. Chief among them is the ability to put aside personal preferences about teaching and "see" through the lens of the rubric. Other important characteristics are:

- **Expertise.** Good candidates are knowledgeable about the teaching to be observed, and they understand the rubric and its use in feedback and evaluation. A good place to find such candidates is among those already certified to observe, if they exist.
- **Focus.** Master coders must locate strong and sufficient evidence to support scores and rationales while observing many people interacting in an imperfect video.
- **Flexibility.** Master coders must defend their own views, but also be open-minded, be willing to hear and consider the views of others, and be convinced to change when the evidence warrants.
- **Patience.** Videos often must be watched over and over in the process of gathering evidence, assigning scores, recording timestamps, and resolving the details of the final rationale.

MASTER CODER TRAINING

Even for the most experienced observers, master coding takes time to learn. Live observation of teachers for feedback and evaluation is different than mining raw video for segments that illustrate teaching components at particular levels of performance. Few observers have had to defend their decisions to others who have formed their own judgments of the same instruction.

It can take three days of training before a group of coders can see and say the same things—an indication that they are ready to master code. Quality training requires:

- **A skilled facilitator.** Whoever trains master coders should possess the same characteristics listed above for the coders themselves. In addition, they must be able to guide a group through questioning and to resolve disagreement not by executive decision, but through discussion grounded in the rubric. This person also may serve as, or work closely with, a process manager who coordinates the assignment of videos and collection of codes once training is complete.
- **Video of teaching.** In advance of the training, organizers need to identify a set of videos that can be used to walk trainees through the coding process and allow them to practice. Discussion and practice scoring of six lesson-length videos can easily consume two to three days of training. These videos should be of sufficient audio and visual quality to code, and they should reflect authentic teaching practice (not test preparation or a guest speaker) and a range of practices, levels of teaching quality, grades, subjects, and demographics (among both teachers and students).

- **Review of the rubric.** Before attempting to code, participants should have a thorough grasp of the observation instrument's rationale, organization, and language. As a group, they should have the opportunity to ask and discuss clarifying questions about how each of a rubric's components is described.

- **Tools and protocols.** Master coding involves the detailed annotation of video. Relevant observed behavior is recorded, these notes are organized by the teaching components represented, and then they are matched to the most appropriate levels of performance based on the rubric's descriptions. Whether on paper or computer, templates are needed to facilitate the process; training should build familiarity with their use.

- **Modeling and practice.** Initially, coders should be led through the process slowly, with frequent pauses to discuss the challenges and importance of each step. They should get several opportunities to practice in a low-stakes environment under the guidance of their trainer and with feedback from their fellow coders. The release to independent scoring should be gradual.

The general approach of master coding is not new. Examples of student writing that have been scored and annotated by expert reviewers have long been used to train assessors of essays that are part of standardized tests. Master coding also is used to train evaluators who assess practice based on teacher-provided artifacts, such as lesson plans and student assignments. Evaluators need examples to anchor their judgment. The full brief focuses narrowly on master coding in the context of a particular medium—video—but one that has become especially relevant amid recent advances in technology and the increased prominence of classroom observation as a tool to promote teacher effectiveness.

Boston Public Schools teachers John Cheesman and Maya Smith participate in a master coding "boot camp" training.

"After taking part in master coding I now have a picture in my mind of what the rubrics are trying to say. When I observe I find that I'm looking for evidence and matching evidence to the rubrics more smoothly, more quickly. I also give more specific feedback. Instead of arguing about whether the evidence is effective or not effective, the conversation is, 'I see where you're coming from: it's developing. How can I be effective?' "

Keith Remillard, Principal, Wakefield Hills Elementary School; West Warwick, RI

BILL & MELINDA GATES *foundation*

www.gatesfoundation.org

To download the full brief, plus other Measures of Effective Teaching (MET) reports, go to www.metproject.org.

ABOUT THE MET PROJECT: Completed in 2012, the MET project was a three-year research partnership of academics, teachers, and education organizations committed to investigating better ways to identify and develop effective teaching. Funding came from the Bill & Melinda Gates Foundation.