Florida's K – 3 Mathematics Formative Assessment System

Developing a Valid and Reliable Observational Measure of Formative Assessment

Laura Lang, Principal Investigator

Robert Schoen, Co-Investigator

Mark LaVenia, Methodologist

Maureen Oberlin, Project Manager

2013 NCTM Research Pre-session, Denver, CO

Overview of Presentation

- Background
- Research Question
- Methods
  - Sample and observation design
  - Observer training
  - Instrumentation
  - Analytic strategy
- Results & Conclusions
- Next Steps

Florida's Definition of Formative Assessment

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes.

(McManus, 2008, p. 3)

Formative Assessment

Formative assessment elicits students' thinking, helping teachers to:

- Diagnose critical misconceptions and errors.
- Identify students' progress toward standard mastery.
- **Differentiate instruction** based on students' understanding (i.e., individualize interventional strategies and regroup students for optimal learning opportunities).
- Provide students the deep understanding needed for future mathematics learning.

Florida's Math Formative Assessment System (MFAS)

- MFAS is a freely available, web-based electronic performance support system that provides mathematics tasks, rubrics, and professional development modules for teachers in grades K–8.
- To date, three randomized field trials have been conducted to study the effects of using MFAS.
  - Studies have varied duration and scope as well as levels of teacher support.

MFAS Characteristics

- MFAS is focused on formative assessment occurring day-to-day and minute-to-minute
- MFAS design leaves teachers to decide whether tasks are implemented in one-on-one interviews, small groups, or whole class
- MFAS supports ongoing teacher collaboration around student thinking and instruction
- MFAS is designed to support differentiated instruction

MFAS Efficacy Trials

Three randomized controlled trials have been conducted on MFAS:

1. a nine-week pilot study in 2010 with K–3 teacher teams using tasks aligned with Florida's NGSSS;
2. a semester-long pilot study in 2012 with grade 2–3 teacher teams using tasks aligned with Florida's NGSSS; and
3. a year-long study currently underway in 2012-13 with K–1 teacher teams using tasks aligned with the CCSS.

Research Question

Does our classroom observation instrument demonstrate desirable psychometric properties?

Instrumentation

- Observation Protocol for Formative Assessment in Mathematics (OPFAM; 2012)
  - intended for use in assessing the degree to which teacher practice associated with formative assessment and differentiated instruction was occurring in the classroom
  - 34 items with Likert-type response categories ranging from 1 (*Not at all Descriptive*) to 4 (*Highly descriptive*)
  - Also included 18 items with nominal (*Yes*, *No*, *N/A or Whole-group*, *Small-group*, *One-on-one*) response categories and 2 open-ended comment fields

Six Subscales

Subscales were aligned to Wiliam and Thompson's (2008) and Wiliam's (2010) formative assessment strategies:

1. clarifying and sharing learning intentions and criteria for success
2. engineering effective classroom discussions and tasks that elicit evidence of learning
3. providing feedback that moves learners forward
4. activating students as instructional resources for one another
5. activating students as the owners of their own learning
6. adjusting the instructional plan based on formative assessment results when the evidence of learning indicates it is warranted (The Big Idea)

Sample and Observation Design

- 2012 semester-long pilot study sample
  - 21 schools across three districts
    - 10 business-as-usual control schools and 11 treatment schools
  - 160 grade 2 and grade 3 teachers
- Classroom observation design
  - 10 (of the 21) schools were selected for classroom observation
    - 1 treatment and 1 control school from each quintile of %FRL
  - 72 classrooms were observed by trained observers
    - 43 (60%) of the classrooms were observed by a single observer and 29 (40%) of the classrooms were observed by a pair of observers
    - 101 classroom observations were distributed among 10 trained observers
  - Observation window spanned a 2-week period April 2012

Observer Training

- Observer training
  - One half-day pre-seminar preparation
    - Completion of online PD module on math formative assessment
    - Independent study of the 23 page observation protocol training manual
  - 2-day seminar/practicum
    - Day 1: Seminar on key terminology and attributes of formative assessment; rating a pre-calibrated video lesson, and discussion on rating divergence; rating of two live classrooms
    - Day 2: Discussion on rating divergence; rating of another live classroom and discussion of rating divergence
  - Mid-observation window review
    - Trainer paired with each observer during classroom observation; held post-observation conference to discuss rating divergence

Analytic Strategy

- Confirmatory Factor Analysis (CFA) to investigate model fit
- Generalizability Theory (G-theory) coefficients to investigate subscale internal consistency and inter-rater reliability
- Multi-level regression to investigate predictive validity of classroom practice on student mathematics performance

Model Fit

- All items except one had an $R^2 > .51$
- All six factors had statistically significant inter-correlation
- Goodness-of-fit indices were as follows: χ^2/df = 603.78/512 = 1.18; RMSEA = 0.05; CFI and TLI = .98; and WRMR = .79
- Thus, even with one low-reliability item and a possible benefit of collapsing two of the subscales, model fit surpassed criteria for close fit on all indices

Subscale Reliability Coefficients

Subscale	Internal consistency	Inter-rater reliability
1. Clarifying and sharing learning intentions and criteria for success	.92	.69
2. Engineering classroom discussion and tasks that elicit evidence of learning	.84	.49
3. Providing feedback that moves learners forward	.92	.70
4. Activating students as owners of their own learning	.96	.87
5. Activating students as instructional resources for one another	.89	.63
6. Adjusting the instructional plan based on formative assessment results	.69	.57

Note. N cases = 72; N observations = 101. Subscales have 6, 7, 7, 7, 3, and 4 items, respectively.

Exploratory Analysis

- Model respecification using 19 items with acceptable IRR
- Subscales 2, 3, and 6 collapsed into a single factor
- Subscales 4 and 5 collapsed into a single factor

Predictive Validity: Relation between Classroom Practice and Grade 2 Student Achievement

Classroom observation Factors	Estimate (SE)	p-value	Effect Size
Clarifying and sharing learning intentions	0.48 (0.11)	< .001	**0.61**
Eliciting evidence of learning and providing feedback	0.36 (0.13)	.006	**0.45**
Student self- and peer-assessment	0.25 (0.15)	.089	0.29

Results & Conclusions

- Analyses suggested the structure of the protocol was valid with internally consistent subscales
  - However, inter-rater reliability analyses suggested sub-optimal agreement between observers
- Predictive validity for target classroom practices on student math performance underscores this as a promising area of investigation
- Observed psychometric properties suggest revision to some of the items as well as to the training manual and sessions is warranted

Next Steps

- The revised protocol was implemented April 1 to 12, 2013 in 31 schools in 3 Florida school districts
  – 124 teachers were scheduled to be observed; half of them by a pair of observers
  – 19 trained observers were employed
- Analyses are underway
- The revised protocol (OPFAM, 2013) training manual can be found at:
  – https://www.dropbox.com/s/g8ed7ra9aelzrrt/FCR-STEM_OPFAM_TRAINING%20MANUAL_4-11-2013.pdf

Increasing Inter-rater Reliability

Some Sources of Bias

- Leniency Bias
- Severity Bias
- Central Tendency Bias
- Halo Bias
- Horns Bias
- Contrast Bias

Classroom in Video

- Kindergarten class
- Spring 2009
- Culminating problem in a unit on composing 7

The 2011 – 2013 MFAS-CCSS Team

- Laura Lang, Principal Investigator
- Robert Schoen, Co-Principal Investigator
- Maureen Oberlin, Project Manager
- Michael Anderson (Pinellas)
- Charity Bauduin (Okaloosa)
- Zachary Champagne (Duval)
- Michelle Robinson (Seminole)
- Linda Walker (Walton)
- Andy Reeves (Pinellas)
- Mark LaVenia, Methodologist
- Anabelle Andon, ELL Support

Questions and Contact information

Laura Lang, Principal Investigator

llang@lsi.fsu.edu

Robert Schoen, Co-principal Investigator

rschoen@lsi.fsu.edu

Revised OPFAM (2013): Subscale 1

Clarifying and sharing learning intentions and criteria for success						
1. Teacher communicates *learning goal*(s) to students.	4	3	2	1	0	N/A
2. Teacher communicates *success criteria* to students.	4	3	2	1	0	N/A
3. Teacher refers to *success criteria* during the lesson.	4	3	2	1	0	N/A
4. *Success criteria* are aligned to *learning goal*(s).	4	3	2	1	0	N/A
5. *Success criteria* relate to what students will say, do, make or write to show *evidence* of learning.	4	3	2	1	0	N/A
6. The enacted lesson aligns with the *learning goal*(s).	4	3	2	1	0	N/A

Revised OPFAM (2013): Example rubric

1. Teacher communicates *learning goal*(s) to students.	**4**	**3**	**2**	**1**	**0**	**N/A**

Rubric 1: Communicates Learning Goal	
Note on High Rating	A factor possibly contributing to a high rating on this item is the learning goal[a] is clearly explained and written in language students can clearly understand. Rating of communication clarity should not be based solely on what is communicated at the outset of the lesson, but should allow for the teacher to roll-out the explanation of learning goal(s)[a] as the lesson progresses.
4	Learning goal[a] is clearly explained to students.
3	Learning goal[a] is mostly explained to students.
2	Learning goal[a] is somewhat explained to students.
1	Learning goal[a] is not communicated to the students (neither verbally stated nor in writing).
0	There does not appear to be a mathematics learning goal[a] intended for the activity.
Note on Low Rating	A factor possibly contributing to a low rating on this item is the learning goal[a] is not written, even though it clearly would be appropriate to do so with these students (e.g., are of reading age, without visual impairment).

Revised OPFAM (2013): Subscale 2

Engineering classroom discussion and tasks that elicit evidence of learning						
7. Teacher presents tasks that promote student mathematical analysis.	4	3	2	1	0	N/A
8. Teacher observes students in the practice of doing mathematics and listens to their mathematics conversations.	4	3	2	1	0	N/A
9. Teacher poses problems and prompts students to share their thinking about the mathematics and how they are approaching the problem.	4	3	2	1	0	N/A
10. Teacher uses wait-time to provide adequate time for cognitive processing.	4	3	2	1	0	N/A
11. Teacher follows up student responses by eliciting student explanations and reasoning.	4	3	2	1	0	N/A
12. Teacher ensures that the student understands the problem.	4	3	2	1	0	N/A
13. Teacher explores what the student has already done.	4	3	2	1	0	N/A
14. Teacher uses revoicing strategies.	4	3	2	1	0	N/A

Revised OPFAM (2013): Subscale 3

Providing feedback that moves learners forward						
15. Teacher feedback provides suggestions to students about what they can do to progress from their current learning status toward the desired *learning goal*.	**4**	**3**	**2**	**1**	**0**	**N/A**
16. Teacher feedback is limited to manageable units.	**4**	**3**	**2**	**1**	**0**	**N/A**
17. Teacher feedback attends to details in the student's reasoning, strategy, or algorithm.	**4**	**3**	**2**	**1**	**0**	**N/A**
18. Teacher feedback to students emphasizes effort.	**4**	**3**	**2**	**1**	**0**	**N/A**
19. Teacher feedback turns learner mistakes into learning opportunities.	**4**	**3**	**2**	**1**	**0**	**N/A**
20. Teacher reminds the student to use other strategies.	**4**	**3**	**2**	**1**	**0**	**N/A**
21. Teacher feedback is presented in more than one modality (e.g., text, visual/graphic, verbal).	**4**	**3**	**2**	**1**	**0**	**N/A**

Revised OPFAM (2013): Subscale 4

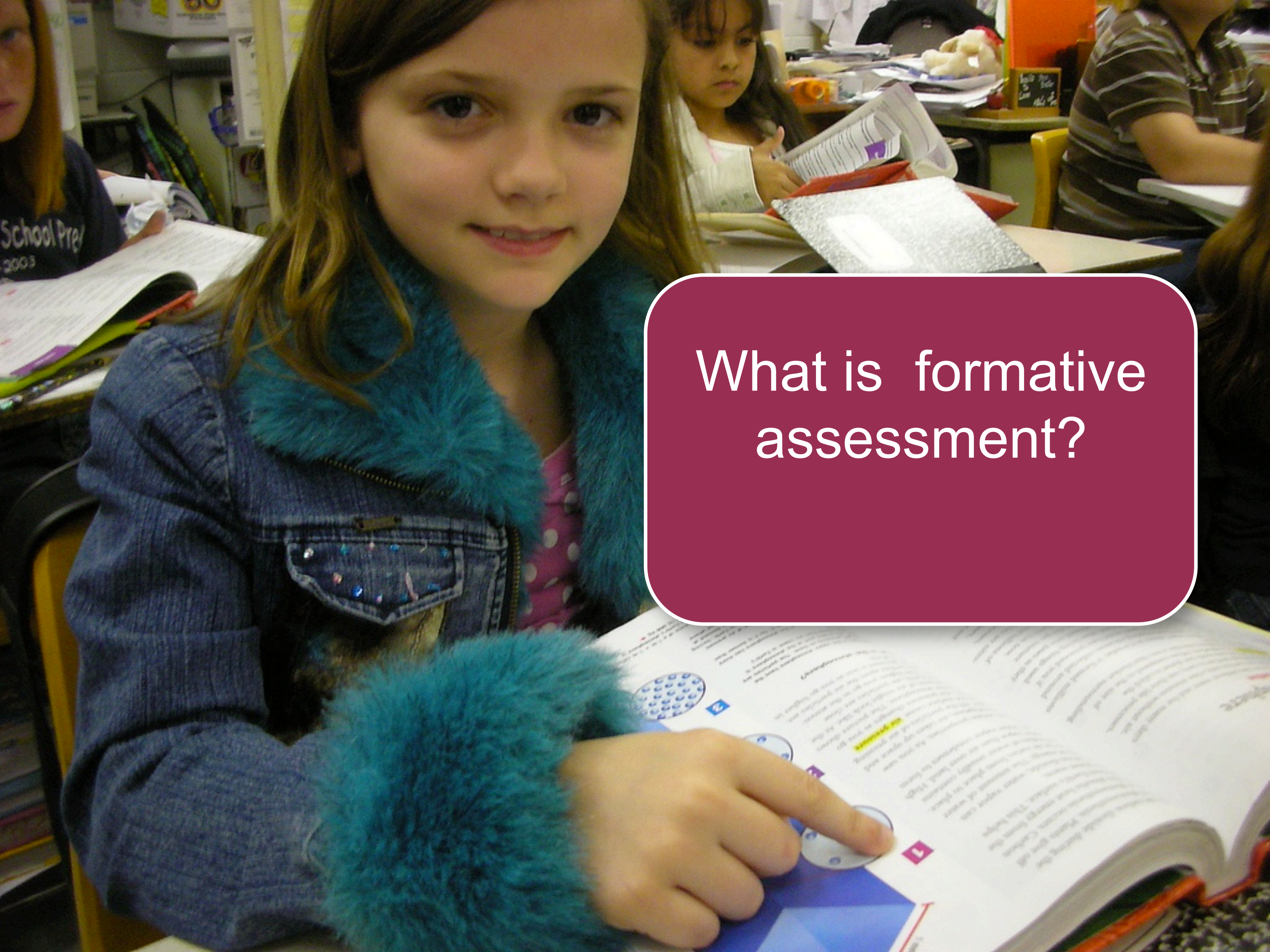
Activating students as instructional resources for one another						
22. Teacher facilitates the sharing of students' thinking to contribute to group talk and help peers.	4	3	2	1	0	N/A
23. Teacher provides opportunities for students to explain their thinking to other students and think about other students' reasoning, strategies, or algorithms.	4	3	2	1	0	N/A
24. Teacher provides opportunities for students to give elaborated peer feedback.	4	3	2	1	0	N/A
25. Teacher provides opportunities for students to use peer feedback to improve their learning.	4	3	2	1	0	N/A
26. Students' contributions link to and build on each other.	4	3	2	1	0	N/A

Revised OPFAM (2013): Subscale 5

Activating students as the owners of their own learning						
27. Teacher provides a system that encourages students to monitor their own learning in relation to the *learning goal*.	**4**	**3**	**2**	**1**	**0**	**N/A**
28. Teacher promotes student reflection on the reasoning, strategy, or algorithm the student just used.	**4**	**3**	**2**	**1**	**0**	**N/A**
29. Students demonstrate productive engagement with mathematics.	**4**	**3**	**2**	**1**	**0**	**N/A**
30. Students reflect on and monitor their learning in relation to the *learning goal*.	**4**	**3**	**2**	**1**	**0**	**N/A**

Revised OPFAM (2013): Subscale 6

Adjusting the instructional plan based on formative assessment results						
Support student thinking before a correct answer is given:						
31. Teacher provides increased support for students who have the lowest level of knowledge in relation to the *learning goal*.	**4**	**3**	**2**	**1**	**0**	**N/A**
32. Teacher changes the mathematics in the problem to match the student's level of understanding.	**4**	**3**	**2**	**1**	**0**	**N/A**
33. Teacher provides linguistic scaffolding and supports cultural congruence, where appropriate.	**4**	**3**	**2**	**1**	**0**	**N/A**
Extends student thinking after a correct answer is given:						
34. Teacher encourages the student to explore multiple strategies and their connections.	**4**	**3**	**2**	**1**	**0**	**N/A**
35. Teacher connects the student's thinking to symbolic notation.	**4**	**3**	**2**	**1**	**0**	**N/A**
36. Teacher generates follow-up problems linked to the problem the student just completed.	**4**	**3**	**2**	**1**	**0**	**N/A**


What is formative assessment?