

Embedding formative assessment with teacher learning communities

Dylan Wiliam (@dylanwiliam)

www.dylanwiliam.net

Overview: Science and Design

2

- We need to improve student achievement
- This requires improving teacher quality
- Improving the quality of entrants takes too long
- So we have to make the teachers we have better **Science**
- We can change teachers in a range of ways
- Some will benefit students, and some will not
- Those that do involve changes in teacher practice

- Changing practice requires new kinds of teacher learning **Design**
- And new models of professional development



Raising achievement matters

3

- For individuals:
 - ▣ Increased lifetime salary
 - ▣ Improved health
 - ▣ Longer life
- For society:
 - ▣ Lower criminal justice costs
 - ▣ Lower healthcare costs
 - ▣ Increased economic growth:
 - Net present value to Sweden of a 25-point increase on PISA: SEK 13 trillion (Hanushek & Woessman, 2015)
 - Net present value to Sweden of getting all students to 420 on PISA: SEK 8 trillion



Formative assessment: what it is and what it isn't; when it works and when it doesn't

Dylan Wiliam

Relevant studies

5

- Fuchs & Fuchs (1986)
- Natriello (1987)
- Crooks (1988)
- Bangert-Drowns et al. (1991)
- Dempster (1991, 1992)
- Elshout-Mohr (1994)
- Kluger & DeNisi (1996)
- Black & Wiliam (1998)
- Nyquist (2003)
- Brookhart (2004)
- Allal & Lopez (2005)
- Köller (2005)
- Brookhart (2007)
- Wiliam (2007)
- Hattie & Timperley (2007)
- Shute (2008)



The formative assessment hijack

6

- Long-cycle:
 - ▣ Span: across units, terms
 - ▣ Length: four weeks to one year
 - ▣ Impact: student monitoring; curriculum alignment
- Medium-cycle:
 - ▣ Span: within and between teaching units
 - ▣ Length: one to four weeks
 - ▣ Impact: improved, student-involved assessment; teacher cognition about learning
- Short-cycle:
 - ▣ Span: within and between lessons
 - ▣ Length:
 - day-by-day: 24 to 48 hours
 - minute-by-minute: five seconds to two hours
 - ▣ Impact: classroom practice; student engagement



Unpacking formative assessment

7

	Where the learner is going	Where the learner is	How to get there
Teacher	Clarifying, sharing and understanding learning intentions	Engineering effective discussions, tasks, and activities that elicit evidence of learning	Providing feedback that moves learners forward
Peer		Activating students as learning resources for one another	
Learner		Activating students as owners of their own learning	

And one big idea

8

	Where the learner is going	Where the learner is	How to get there
Teacher	<p>Using evidence of achievement to adapt what happens in classrooms to meet learner needs</p>		
Peer			
Learner			

An educational positioning system

9

- A good teacher:
 - ▣ Establishes where the students are in their learning
 - ▣ Identifies the learning destination
 - ▣ Carefully plans a route
 - ▣ Begins the learning journey
 - ▣ Makes regular checks on progress on the way
 - ▣ Makes adjustments to the course as conditions dictate



Strategies and practical techniques for classroom formative assessment

Clarifying, sharing and understanding learning intentions

Sharing learning intentions

- 3 teachers each teaching 4 7th grade science classes in two US schools
- 14 week experiment
- 7 two-week projects, each scored 2-10
- All teaching the same, except:
- For a part of each week
 - Two of each teacher's classes discusses their likes and dislikes about the teaching (control)
 - The other two classes discuss how their work will be assessed



Sharing learning intentions

13

	Comprehensive Test of Basic Skills		
Group	Low	Middle	High
Likes and dislikes	4.6	5.9	6.6
Reflective assessment			

Who benefits most from reflective assessment?

1. Low achievers
2. Average students
3. High achievers
4. All students benefit equally



Share learning intentions

14

- Explain learning intentions at start of lesson/unit:
 - ▣ Learning intentions
 - ▣ Success criteria
- Consider providing learning intentions and success criteria in students' language
- Use posters of key words to talk about learning:
 - ▣ E.g., describe, explain, evaluate
- Use planning and writing frames judiciously
- Use annotated examples of different standards to “flesh out” assessment rubrics (e.g., lab reports).
- Provide opportunities for students to design their own tests.



Engineering effective discussions,
activities, and classroom tasks that elicit
evidence of learning

Kinds of questions: Israel

Which fraction is the smallest? a) $\frac{1}{6}$, b) $\frac{2}{3}$, c) $\frac{1}{3}$, d) $\frac{1}{2}$.

Success rate 88%

Which fraction is the largest? a) $\frac{4}{5}$, b) $\frac{3}{4}$, c) $\frac{5}{8}$, d) $\frac{7}{10}$.

Success rate 46%; 39% chose (b)

Vinner (1997)



Eliciting evidence

- Key idea: questioning should
 - ▣ cause thinking
 - ▣ provide data that informs teaching
- Improving teacher questioning
 - ▣ generating questions with colleagues
 - ▣ low-order vs. high-order not closed vs. open
 - ▣ appropriate wait-time
- Getting away from I-R-E
 - ▣ basketball rather than serial table-tennis
 - ▣ ‘No hands up’ (except to ask a question)
 - ▣ ‘Hot Seat’ questioning
- All-student response systems
 - ▣ ABCD cards, “show-me” boards, exit passes



Eliciting evidence: Kinds of questions

Questioning in science: Discussion

Ice-cubes are added to a glass of water. What happens to the level of the water as the ice-cubes melt?

- A. The level of the water drops
- B. The level of the water stays the same
- C. The level of the water increases
- D. You need more information to be sure



Questioning in science: Diagnosis (2)

20

Version 1

Which of these are living?

- A. Rock
- B. Cat
- C. Table
- D. Bird

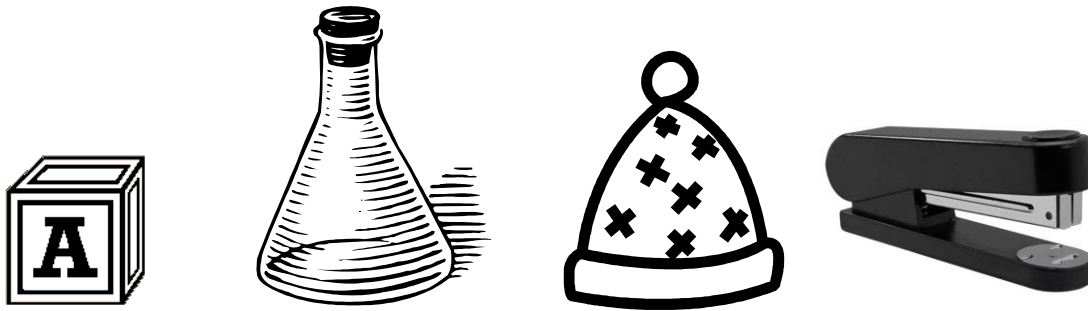
Version 2

Which of these are living?

- A. Grass
- B. Bus
- C. Computer
- D. Tree



Questioning in science: Diagnosis (3)



Sheena leaves a wooden block, a glass flask, a woolly hat, and a metal stapler on a table overnight. What can she say about their temperatures the next morning?

- A. The stapler will be colder than the other objects
- B. The woolly hat will be warmer than the other objects
- C. The temperatures of all four objects will be different
- D. The temperatures of all four objects will be the same



Questioning in mathematics: Diagnosis (2)

What is the median for the following data set?

38 74 22 44 96 22 19 53

- A. 22
- B. 38 and 44
- C. 41
- D. 46
- E. 77
- F. This data set has no median



Questioning in mathematics: Diagnosis (3)

What can you say about the means of the following two data sets?

Set 1: 10 12 13 15

Set 2: 10 12 13 15 0

- A. The two sets have the same mean.
- B. The two sets have different means.
- C. It depends on whether you choose to count the zero.



Questioning in English: Diagnosis (1)

Where is the verb in this sentence?

The dog ran across the road

↑ ↑ ↑ ↑
A B C D



Questioning in English: Diagnosis (2)

Which of these is correct?

- A. Its on its way.
- B. It's on its way.
- C. Its on it's way.
- D. It's on it's way.



Questioning in history: Diagnosis (1)

Why are historians concerned with bias when analyzing sources?

- A. People can never be trusted to tell the truth
- B. People deliberately leave out important details
- C. People are only able to provide meaningful information if they experienced an event firsthand
- D. People interpret the same event in different ways, according to their experience
- E. People are unaware of the motivations for their actions
- F. People get confused about sequences of events



Questioning in MFL: Diagnosis

Which of the following is the correct translation for “I give the book to him”?

- A. Yo lo doy el libro.
- B. Yo doy le el libro.
- C. Yo le doy el libro.
- D. Yo doy lo el libro.
- E. Yo doy el libro le.
- F. Yo doy el libro lo.



Hinge questions

- A hinge question is based on the important concept in a lesson that is critical for students to understand before you move on in the lesson.
- The question should fall about midway during the lesson.
- Every student must respond to the question within two minutes.
- You must be able to collect and interpret the responses from all students in 30 seconds



Providing feedback that moves
learners forward



Kinds of feedback: Israel

- 264 low and high ability grade 6 students in 12 classes in 4 schools; analysis of 132 students at top and bottom of each class
- Same teaching, same aims, same teachers, same classwork
- Three kinds of feedback: grades, comments, grades+comments

	Achievement	Attitude
Grades	no gain	High scorers: positive Low scorers: negative
Comments	30% gain	High scorers: positive Low scorers: positive

Butler (1988)



Responses

	Achievement	Attitude
Grades	no gain	High scorers: positive Low scorers: negative
Comments	30% gain	High scorers: positive Low scorers: positive

What happened for students given both grades and comments?

- A. Gain: 30%; Attitude: all positive
- B. Gain: 30%; Attitude: high scorers positive, low scorers negative
- C. Gain: 0%; Attitude: all positive
- D. Gain: 0%; Attitude: high scorers positive, low scorers negative
- E. Something else



Åsas exempel

32

”Jag har kollat igenom kommentarerna men eftersom betyget står på så blir man ändå smått blind av det och fokuserar lite för mycket på det. Så personligen (även om jag mycket möjligt skulle kunna klaga om jag inte får betyget skrivet) skulle vilja att du inte gör det för att jag märker att jag lägger mer fokus på kommentarerna och utvecklas bättre då betyget inte står på papperet.”



Kinds of feedback: Israel (2)

- 200 grade 5 and 6 Israeli students
- Divergent thinking tasks
- 4 matched groups
 - ▣ experimental group 1 (EG1); comments
 - ▣ experimental group 2 (EG2); grades
 - ▣ experimental group 3 (EG3); praise
 - ▣ control group (CG); no feedback
- Achievement
 - ▣ $EG1 > (EG2 \approx EG3 \approx CG)$
- Ego-involvement
 - ▣ $(EG2 \approx EG3) > (EG1 \approx CG)$



Effects of feedback

- Kluger & DeNisi (1996) review of 3000 research reports
- Excluding those:
 - ▣ without adequate controls
 - ▣ with poor design
 - ▣ with fewer than 10 participants
 - ▣ where performance was not measured
 - ▣ without details of effect sizes
- left 131 reports, 607 effect sizes, involving 12652 individuals

- On average, feedback increases achievement
 - ▣ Effect sizes highly variable
 - ▣ 38% (50 out of 131) of effect sizes were negative



Discussion question

35

- How can feedback lower student achievement?



Getting feedback right is hard

Response type	Feedback indicates performance...	
	falls short of goal	exceeds goal
Change behavior	Increase effort	Exert less effort
Change goal	Reduce aspiration	Increase aspiration
Abandon goal	Decide goal is too hard	Decide goal is too easy
Reject feedback	Feedback is ignored	Feedback is ignored



Provide feedback that moves learning on

- Key idea: feedback should:
 - ▣ Cause thinking
 - ▣ Provide guidance on how to improve
- Comment-only marking
- Focused marking
- Explicit reference to mark-schemes/rubrics
- Suggestions on how to improve:
 - ▣ Not giving complete solutions
- Re-timing assessment:
 - ▣ E.g., three-fourths-of-the-way-through-a-unit test



Discussion question

38

- What are the obstacles to introducing more effective forms of feedback?



Activating students as learning resources for one another

Cooperative learning: a research success story

40

- Two essential components
 - ▣ Group goals:
 - so students are working *as* a group, not just *in* a group
 - ▣ Individual accountability:
 - the best learning efforts of every member of the group must be necessary for the group to succeed, and
 - the performance of each group member must be clearly visible and quantifiable to the other group members

Slavin, Hurley and Chamberlain (2003)



How does cooperative learning work?

□ Four mechanisms

- Motivation: students help their peers to learn because, in well-structured cooperative learning settings, it is in their own interests to do so, and so effort is increased;
- Social cohesion: students help their peers because they care about the group, again leading to increased effort;
- Personalization: students learn more because more able peers can engage with the particular difficulties a student is having;
- Cognitive elaboration: those who provide help in group settings are forced to think through the ideas more clearly.

Slavin, Hurley and Chamberlain (2003)



Help students be learning resources

- Students assessing their peers' work:
 - ▣ “Pre-flight checklist”
 - ▣ “Two stars and a wish”
 - ▣ Choose-swap-choose
 - ▣ Daily sign-in
- Training students to pose questions/identifying group weaknesses
- End-of-lesson students' review



Activating students as owners of their own learning

Help students own their own learning

- Students assessing their own work:
 - ▣ With rubrics
 - ▣ With exemplars
- Self-assessment of understanding:
 - ▣ Learning portfolio
 - ▣ Traffic lights
 - ▣ Red/green discs
 - ▣ Coloured cups
 - ▣ Plus/minus/interesting



- "
- +
- "
- I get that ball park estimates are supposed to be simple. Meghan
 - I know that you have to look at it and say ^{Frankie} "Ohh?"
 - I know when I am adding the number I end up with must be bigger then the one I started at. Jon
 - I get most of the problems. Julianna.
 - It was ~~in~~ easy for me because on the first one it says 328 and I took the # 2. and I made it a 12. Kelly
 - I know that we would have to regroup. Alana
 - ~~it is~~ I know how to do Plus and minus ~~because~~ because we have been doing it ^{for a long time} for a long time.
 - I think because for 4 some years we've been I think I finally know that adding is combining the two numbers in the problem.
 - I think I am good at the partial sums method. ^{TR. Elop}
 - I get it when you cross out a number and make it a new one. Emma
 - I know when you can't - from both columns you go to the third column and take that from it. Olivia

I know when my answer is right the ball park estimate is close the the answer. Brendan



I am still a tiny bit confused about subtraction regrouping. Meaghan

I am a little bit confused about ball park estimate. Julianne

I get confused because sometimes

I don't get the problem. Frankie

I am confused when you subtract really big numbers.

Like 1,000 something. Jan. I'm still a little bit confused about

regrouping. Trevor

I am confused about a little of the subtraction regrouping. Aidan

I am a little confused about the regrouping still. Kelly

Minus is confusing because when you have to regroup twice. Alana

Minus is a little bit hard when you have to

regroup. Darci

I don't understand when you borrow which column to borrow from when both are 0. Olivia.

I am still confused about showing what I did to solve the problem. Brendan

I am a little confused about when you need to subtract. Emma



interesting

Carrying the number over to the next number

Julianne

It's interesting how some people go to the nearest hundred, while others go to the nearest ten. Meaghan

It's interesting how some have to regroup twice. ^{Alana}

It is interesting sometimes how you have to regroup ~~that's~~ Darci

• It's pretty interesting about how you have to really work hard. Frankie

• ~~I am~~ I am interested in borrowing because I didn't just get it yet. I want to really get to know it. Jon

• I find it weird that you could just keep going from column to column when you need to borrow. Olivia

• On the ballpark estimate it is ~~pretty good~~ easy but sometimes confusing. Kelly

• I really think that regrouping is pretty amazing.

• It is cool how addition and subtraction regrouping is just moving numbers and you could get it right easily.



+/-/interesting: responses for “+”

- I got that ball-park estimates are supposed to be simple
- I know that you have to look at it and say “OK”
- I know that when I am adding the number I end up with must be bigger than the one I started at
- I get most of the problems
- It was easy for me because on the first one it says 328 so I took the 2 and made it a 12
- I know that we would have to regroup
- I know how to do plus and minus because we have been doing it for a long time
- I get it when you cross out a number and make it a new one
- I know that when you can't – from both colomes you go to the third colome and take that from it
- I know that when my answer is right the ball park estimate is close to it



+/-/interesting: responses for “-”

- I am still a tiny bit confused about subtraction regrouping
- I am a little bit confused about ball park estimates
- I get confused because sometimes I don't get the problem
- I am confused when you subtract really big numbers like 1,000 something
- I'm still a little bit confused about regrouping
- Minus is confusing when you have to regroup twice
- Minus is a little bit hard when you have to regroup
- I don't understand when you borrow which colome you borrow from when both are 0
- I am still confused about showing what I did to solve the problem
- I am a little confused about when you need to subtract



+/-/interesting: responses for “interesting”

- Carrying the number over to the next number
- It's interesting how some people go to the nearest hundred while some go to the nearest ten
- It's interesting how some have to regroup twice
- It's pretty interesting about how you have to work really hard
- I am interested in borrowing because I didn't just get it yet. I want to really get to know it
- I find it weird that you could just keep going from colome to colome when you need to borrow
- On the ball park estimate it is easy but sometimes hard
- I really think that regrouping is pretty amazing
- It is cool how addition and subtraction regrouping is just moving numbers and you could get it right easily



Self-assessment in the early years

51





All ready for action in year four...

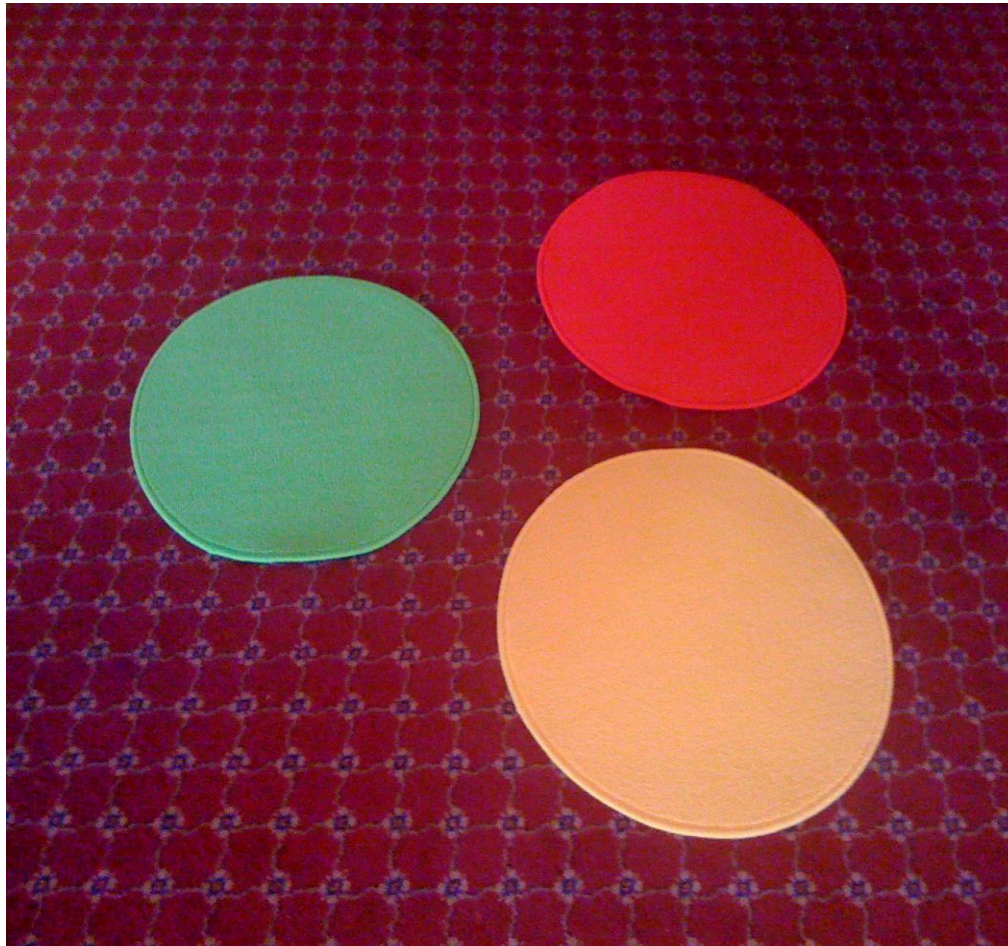
53



Tell me about you...



IKEA mats...



Technique review

So much for the easy bit

A model for teacher learning

58

- Content, then process
- Content (what we want teachers to change):
 - Evidence
 - Ideas (strategies and techniques)**Science**
- Process (how to go about change):
 - Choice
 - Flexibility
 - Small steps
 - Accountability
 - Support**Design**



Choice



A strengths-based approach to change

60

- Belbin inventory (Management teams: Why they succeed or fail):
 - ▣ Eight team roles (defined as “a tendency to behave, contribute and interrelate with others in a particular way”):
 - Company worker; innovator; shaper; chairperson; resource investigator; monitor/evaluator; completer/finisher; team worker
 - ▣ Key ideas:
 - People rarely sustain “out-of-role” behaviour, especially under stress
 - Each role has strengths and allowable weaknesses
- Each teacher’s personal approach to teaching is similar:
 - ▣ Some teachers’ weaknesses require immediate attention
 - ▣ For most, however, students benefit more from the development of teachers’ strengths



Flexibility



Strategies vs. techniques

62

- Distinguish between strategies and techniques:
 - Strategies define the territory of formative assessment (no-brainers)
 - Teachers are responsible for choice of techniques:
 - Allows for customization; caters for local context
 - Creates ownership; shares responsibility
- Key requirements of techniques:
 - They embody the deep cognitive and affective principles that research shows are important
 - They are seen as relevant, feasible and acceptable



Small steps

Why is teacher change so slow?

64

- Because of the nature of teacher expertise
- According to Berliner (1994), experts:
 - ▣ excel mainly in their own domain
 - ▣ develop automaticity for operations needed for their goals
 - ▣ are more sensitive to the task demands and social situations
 - ▣ are more opportunistic and flexible than novices
 - ▣ represent problems in qualitatively different ways than novices
 - ▣ have faster and more accurate pattern recognition capabilities
 - ▣ see richer patterns in the areas of their expertise
 - ▣ begin to solve problems slower but bring richer and more personal sources of information to bear



Knowing more than we can say

65

- Six video extracts of a person delivering cardiopulmonary resuscitation (CPR):
 - ▣ Five of the video extracts feature students
 - ▣ One of the video extracts feature an expert
- Videos shown to three groups:
 - ▣ students, experts, instructors
- Success rate in identifying the expert:
 - ▣ Experts 90%
 - ▣ Students 50%
 - ▣ Instructors 30%



Looking at the wrong knowledge

66

- The most powerful teacher knowledge is not explicit:
 - ▣ That's why telling teachers what to do doesn't work
 - ▣ What we know is more than we can say
 - ▣ And that is why most professional development has been relatively ineffective
- Improving practice involves changing habits, not adding knowledge:
 - ▣ That's why it's hard
 - And the hardest bit is not getting new ideas into people's heads
 - It's getting the old ones out
 - ▣ That's why it takes time
- But it doesn't happen naturally:
 - ▣ If it did, the most experienced teachers would be the most productive, and that's not true (Hanushek & Rivkin, 2006)



Most of what we do is unconscious

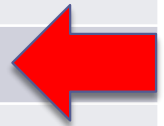
67

Sensory system	Total bandwidth (in bits/second)	Conscious bandwidth (in bits/second)
Eyes	10,000,000	40
Ears	100,000	30
Skin	1,000,000	5
Taste	1,000	1
Smell	100,000	1



Hand hygiene in hospitals

Study	Focus	Compliance rate
Preston, Larson, & Stamm (1981)	Open ward	16%
	ICU	30%
Albert & Condie (1981)	ICU	28% to 41%
Larson (1983)	All wards	45%
Donowitz (1987)	Pediatric ICU	30%
Graham (1990)	ICU	32%
Dubbert (1990)	ICU	81%
Pettinger & Nettleman (1991)	Surgical ICU	51%
Larson, et al. (1992)	Neonatal ICU	29%
Doebbeling, et al. (1992)	ICU	40%
Zimakoff, et al. (1992)	ICU	40%
Meengs, et al. (1994)	ER (Casualty)	32%
Pittet, Mourouga, & Perneger (1999)	All wards	48%
	ICU	36%



Accountability

Making a commitment

70

- Action planning:
 - ▣ Forces teachers to make their ideas concrete and creates a record
 - ▣ Makes the teachers accountable for doing what they promised
 - ▣ Requires each teacher to focus on a small number of changes
 - ▣ Requires the teachers to identify what they will give up or reduce
- A good action plan:
 - ▣ Does not try to change everything at once
 - ▣ Spells out specific changes in teaching practice
 - ▣ Relates to the five “key strategies” of formative assessment
 - ▣ Is achievable within a reasonable period of time
 - ▣ Identifies something that the teacher will no longer do or will do less of



And being held to it

71

“I think specifically what was helpful was the ridiculous NCR [No Carbon Required] forms. I thought that was the dumbest thing, but I’m sitting with my friends and on the NCR form I write down what I am going to do next month.

“Well, it turns out to be a sort of ‘I’m telling my friends I’m going to do this’ and I really actually did it and it was because of that. It was because I wrote it down.

“I was surprised at how strong an incentive that was to do actually do something different...that idea of writing down what you are going to do and then because when they come by the next month you better take out that piece of paper and say ‘Did I do that?’...just the idea of sitting in a group, working out something, and making a commitment...I was impressed about how that actually made me do stuff.”

—Tim, Spruce Central High School



Support

Supportive accountability

73

- What is needed from teachers:
 - ▣ A commitment to:
 - The continual improvement of practice
 - Focus on those things that make a difference to students
- What is needed from leaders:
 - ▣ A commitment to engineer effective learning environments for teachers by:
 - Creating expectations for continually improving practice
 - Keeping the focus on the things that make a difference to students
 - Providing the time, space, dispensation, and support for innovation
 - Supporting risk-taking



Teacher learning communities



- We need to create time and space for teachers to reflect on their practice in a structured way, and to learn from mistakes.

Bransford, Brown & Cocking (1999)

- “Always make new mistakes.”

Esther Dyson

- “Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.”

Beckett (1984)



Teacher learning communities

76

- Plan that the TLC will run for two years
- Identify 10 to 12 interested colleagues:
 - ▣ Conscripts vs. volunteers
 - ▣ Composition:
 - Similar assignments (e.g., early years, math/science)
 - Mixed subject/mixed phase
 - Hybrid
- Secure institutional support for:
 - ▣ Monthly workshops (75–120 minutes each, inside or outside school time)
 - ▣ Time between workshops (two hours per month in school time) for collaborative planning and peer observation
 - ▣ Any necessary waivers from school policies



A “signature pedagogy” for teacher learning

77

- Every monthly TLC workshop should follow the same structure and sequence of activities:
 - ▣ Activity 1: Introduction (5 minutes)
 - ▣ Activity 2: Starter activity (5 minutes)
 - ▣ Activity 3: Feedback (25–50 minutes)
 - ▣ Activity 4: New learning about formative assessment (20–40 minutes)
 - ▣ Activity 5: Personal action planning (15 minutes)
 - ▣ Activity 6: Review of learning (5 minutes)



Every TLC needs a leader

78

- The job of the TLC leader(s):
 - ▣ To ensure that all necessary resources (including refreshments!) are available at workshops
 - ▣ To ensure that the agenda is followed
 - ▣ To maintain a collegial and supportive environment
- But most important of all:
 - ▣ It is not to be the formative assessment “expert”



Peer observation

79

- Run to the agenda of the observed, not the observer:
 - ▣ Observed teacher specifies focus of observation:
 - e.g., teacher wants to increase wait time
 - ▣ Observed teacher specifies what counts as evidence:
 - Provides observer with a stopwatch to log wait times
 - ▣ Observed teacher owns any notes made during the observation



Summary

80

- Raising achievement is important
- Raising achievement requires improving teacher quality
- Improving teacher quality requires teacher professional development
- To be effective, teacher professional development must address:
 - ▣ What teachers do in the classroom
 - ▣ How teachers change what they do in the classroom
- Formative assessment + teacher learning communities:
 - ▣ A point of (uniquely?) high leverage



To find out more...



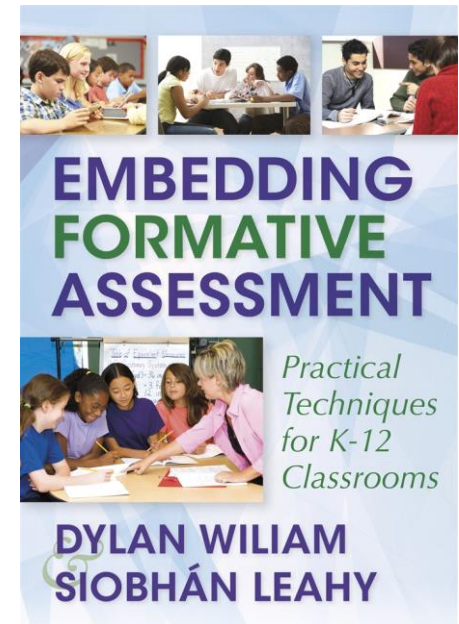
SSAT

inquire
inspire
innovate
impact

Embedding formative assessment

A two-year professional development
pack for schools and colleges:
teacher learning communities in action

Written by
Siobhan Leahy and Dylan Wiliam



www.dylanwiliam.net

