Answering Question 2:
How Will We Know They Have Learned It?

A Solution Tree Event

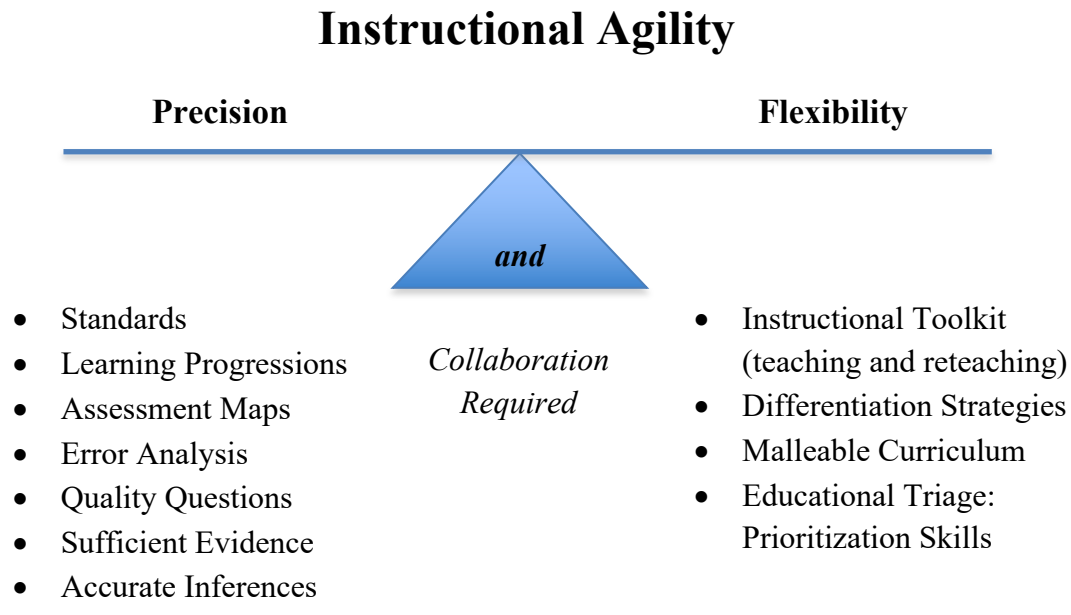
Cassandra Erkens
A PLC and Assessment Associate
casserkens@gmail.com
Twitter: @cerkens #ATAssessment

Assessment practices must build hope, efficacy, and achievement for learners and teachers.

Instructional agility: It is a teacher's intentional maneuver to respond to what is happening as a result of his or her instructional influence. Instructional agility occurs when emerging evidence informs real-time modifications within the context of expected learning.

(Erkens, Schimmer, & Vagle, *Instructional Agility*, 2017)

- **Precision:** Accuracy and exactness
- **Flexibility:** Susceptible to modification or adaptation
- **Emerging evidence:** Student questions, student answers, student conversations, peer feedback, coaching conversations, quizzes, homework, exams, projects, and so on.
- **Real-time modifications:** Extend, enrich, intervene, regroup, correct, acknowledge, question, challenge, validate, prove, explore



The Four I's of Instructional Agility

Instruct	Investigate	Interpret	Intervene
Provide purposeful directions regarding the learners' acquisition of new knowledge, skills, or abilities.	Seek evidence of instructional influence by observing emerging evidence.	Decode the emerging evidence with speed and accuracy for indications of effectiveness, confusion, errors, and so on.	Adjust with responsive instructional maneuvers that keep learning moving forward.

"The vision of practice that underlies the nation's reform agenda requires most teachers to rethink their own practice, to construct new classroom roles and expectations about student outcomes, and to teach in ways they have never taught before."

—Darling-Hammond & McLaughlin, "Policies That Support Professional Development in an Era of Reform," *Phi Delta Kappan, 76*(8): 597–604

Assessment

- The Latin term for assessment—*assidere*—literally means "to walk or sit beside."
- Assessment must be something teachers do with and for learners—not *to* learners.

Assessment: The gathering of clean data with which to make informed decisions (just the facts: mastered 2 of 3 targets, in the remaining target to be mastered, this is the type of error being made….)

Evaluation: The judgment of assessment results (comments made, check marks for completion, totals, percents, or grades)

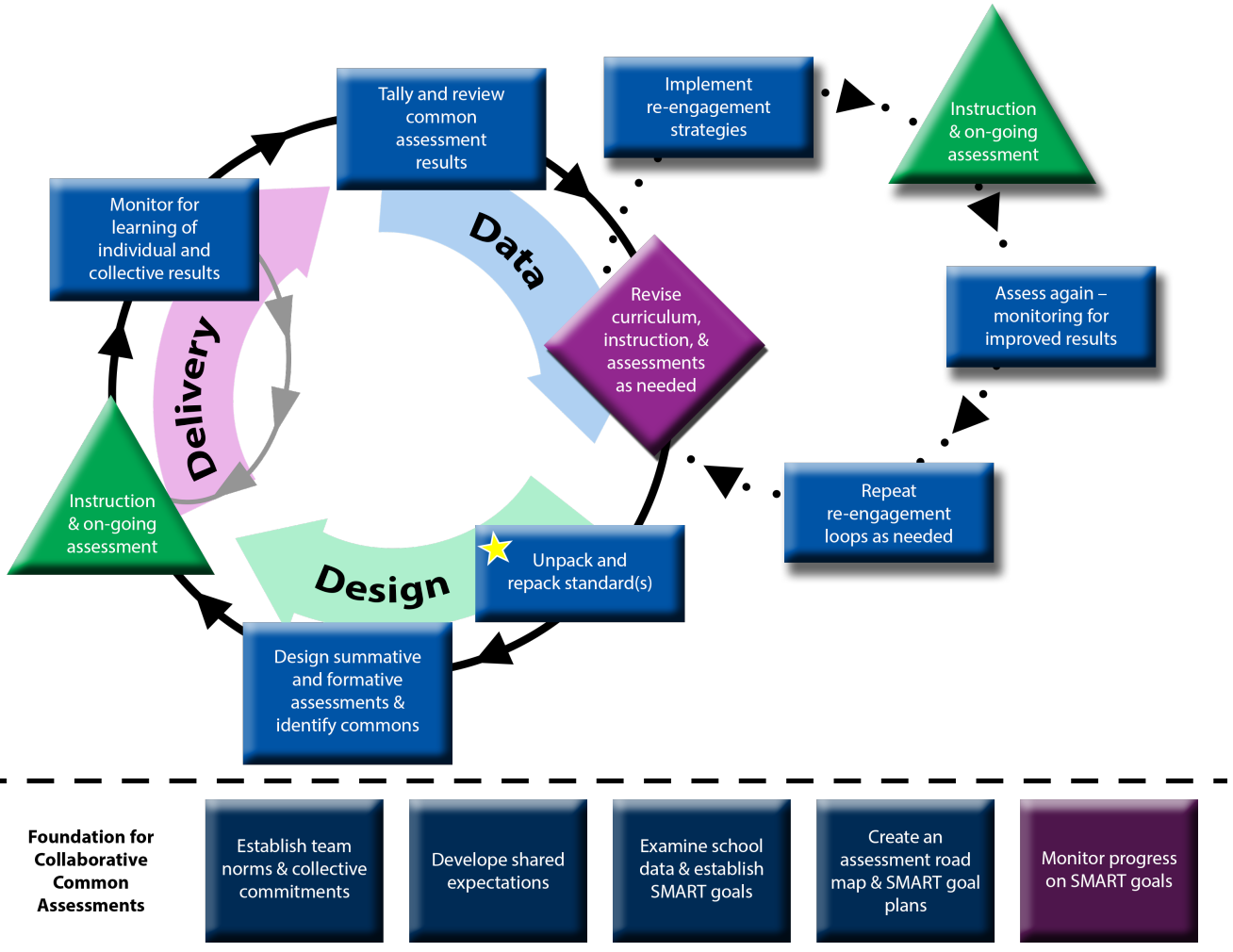
Collaborative Common Assessments

What

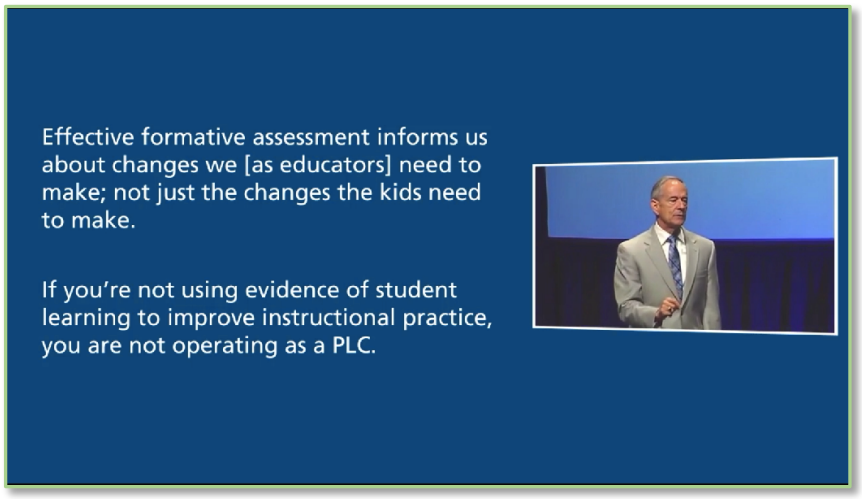
A collaborative common assessment is any assessment, formative or summative, which is *team created or team endorsed* by all of the teachers who share the same standard expectations. The common assessments must be designed *in advance of instruction* and administered in close proximity by all instructors who share a role in administering that assessment. Those who designed or endorsed the assessment must then *collaboratively examine the results* for consistent scoring and shared *instructionally sensitive responses* that address the following:

- Error analysis and appropriate intervention planning for individual learners.
- Curriculum, instruction, and/or assessment modifications.

(Erkens, *Collaborative Common Assessments*, 2016, p. 7)



—Erkens, *The Handbook for Collaborative Common Assessments* (2019), p. 15



How could we use evidence on a meeting-by-meeting basis to better inform our efforts?

Data Moments

An AllThingsPLC.info blog post by Cassandra Erkens, July 26, 2017

Collaborative teams engage in professional learning when they focus on the results of their own efforts. In a professional learning community, data from team-developed common assessments serve as the linchpins of success.

Too often, however, teams are bogged down by data: the data set is too big, the opportunities for gathering the data are too sparse (just one or two common assessments in a quarter), the organization of the data is too time-consuming, the meeting time to discuss the results is too short, etc. For these many reasons, teams often confess they spend more time planning their efforts than examining the results of their efforts. Planning isn't bad; it just isn't sufficient in a professional learning community. Healthy and productive teams always examine the impact of their best-laid plans.

As an alternative, try having data moments in every meeting rather than just awaiting the more formal but infrequent data meetings. Bring small sample sizes of data or even single artifacts of student work to explore together during the first five minutes of your meeting time.

Significant conversations and meaningful next steps can happen during data moments as teams explore questions such as the following:

Moments with single artifacts:

- What score(s) would we give this student based on this work? Do we have inter-rater reliability? If not, what could we do to become more consistent?
- What are the students' strengths? Has the student demonstrated growth over time (e.g. from earlier indications with prior assessments)?
- What are the students' opportunities for growth?
- What instruction have we planned and/or delivered that should have addressed the challenges this student is still facing?
- What could we still do to help this student move forward?
- Do we have the right criteria? The right performance descriptors?
- Is there anything we should change about our assessments and attending tools (rubrics, scales, scoring guides, etc.)?

Data moments with small sample sizes:

- What score(s) would we give these students based on the provided work? Do we have inter-rater reliability?
- What has everyone mastered within the sample size?
- What errors and misconceptions are evident in the sample size?

- Based on our collective expertise, is this sample of evidence representative of the larger group? If so, how do we know? If not, what additional evidence might we need to bring to the table at the next meeting?
- Are there artifacts in the sample set that we could use anonymously in our individual classrooms (use anonymous work from someone else's classroom to avoid exposing your own students) to help students understand strong and weak work?
- How else could we use these few artifacts to help all of our students improve?
- What program improvements (curriculum, instruction, and assessment) do the samples suggest we might need to consider?
- What are plausible next steps for the students represented here, and would those next steps be beneficial to the larger group as well?

It's surprising how much a team can learn after five short minutes of mining data evidence and artifacts. When teams become experienced with data moments, they can be more adept in data meetings. A steady stream of data moments can provide more focus and clarity for a team's planning efforts in the rest of the team meeting.

https://www.allthingsplc.info/blog/view/351/Data+Moments

"My Favorite No" Kindergarten Language Arts (One team's report, New Glarus, WI)

In the spring of the year, the kindergarten team wanted to know if their students could use capital letters and end marks appropriately. They decided to have the students use a 3 x 5 card to write a question for a friend and include the friend's name in the question because the kindergarten students often remember to capitalize their own name but not someone else's. Each teacher had the learners work with their friends to brainstorm a series of interesting questions while sitting on the class carpet. Then they gave the students a 3 x 5 card ask them to write one of their favorite questions to their friend, including his/her name. They gathered the completed cards, and sorted them at the team meeting. They created the following categories to represent their findings:

- Some didn't use end marks at all.
- Some used end marks, but not the right one.
- Some didn't use any capital letters.
- Some used capital letters on the first word of the sentence but not on the proper name.
- A few were making both punctuation and capitalization errors.
- One did not listen to directions and simply wrote "Yes!" in answer to the last question his friend has asked him while brainstorming on the carpet.

From there, the team strategized the instructional responses they would use for each category of error.

"My Favorite Finding" Middle School Social Studies
During a unit on using text evidence and drawing conclusions a middle school social studies teacher gave the learners a passage to read about the law making process. He then gave them a 3 x 5 card and

asked them to draw a conclusion that they could support by using sufficient and valid text evidence from the passage. They were not to put their names on the cards.

After 15 minutes (reading the short passage and writing their conclusions on their cards), the teacher collected the cards. He distributed the criteria for quality conclusions (familiar rubric for the learners), and put the learners into small teams. He asked them to review the criteria again while he sorted the note cards into piles of 1's, 2's, 3's and 4's based on the criteria. Once the cards were sorted, he pulled an example from each pile. He mixed up the order of the 4 cards he selected, and then picked the top card.

He rewrote the card on his Smart Board and asked the teams to examine their rubric so they could score it. They had to generate the appropriate score and be ready to defend it to the rest of the class by

1. backing their score with evidence, and
2. naming the specific strengths and errors within the sample.

At the conclusion, the learners discovered the following errors can happen when trying to draw conclusions:

- Conclusion is simply a restatement of an explicit piece of text evidence.
- Conclusion is based on faulty or inappropriate evidence.
- Conclusion is based on singular or insufficient evidence and misses the complete picture.
- There is sufficient or accurate evidence, but the conclusion is based on weak reasoning or faulty logic.

In the following lessons, the teacher helped the learners begin to strategize ways to fix the specific types of errors.

"My Favorite No" High School Science (One teacher's report, Webster City, Iowa)
During a unit on balancing chemical reactions, a high school science teacher put the following bell-ringer activity on the board:

Balance the equation: $\mathbf{H_2O_2 \rightarrow H_2O + O_2}$

Students entered the room, took 3x5 note cards, and began solving the problem at their desks. Immediately after taking attendance, the teacher gathered the note cards and quickly sorted them into piles for right and wrong answers, making sure no student's name or work was revealed. She selected a card as her "favorite no"—the answer had a few things correct, but also had a common error. To conceal the student's identity, she recreated the "favorite no" in her handwriting. She then asked the class to tell her what was correct with the answer. Finally, the students analyzed what was wrong.

- Correct answer was: $\mathbf{2H_2O_2 \rightarrow 2H_2O + O_2}$
- Her "favorite no" put the coefficients in the wrong place: $\mathbf{H_22O_2 \rightarrow H_22O + O_2}$

Through this process, all students could understand common errors when balancing equations. The teacher found this formative assessment tool to be quick, easy, and supportive for reducing common errors. She noted that students joined in to address errors quickly. With this success, she plans to use the process often.

The Missing Step
Error Analysis

Fisher & Frey (2007): "Misconceptions include preconceived notions, nonscientific beliefs, naïve theories, mixed conceptions, or conceptual misunderstandings" (p. 32).

Identifying, Labeling, and Addressing Errors **Create Stacks Using Student Work**		
	Right and Wrong Answers	**Rubric and Proficiency Scored Answers**
1. Create stacks.	Create two stacks—those that are all right and those with errors.	Create the number of stacks that align with the predetermined proficiency levels.
2. Identify errors (reading error, concept error, reasoning error).	Examine error stack and subdivide into smaller stacks of like errors.	Examine each stack and identify common errors or isolate the components that would move work into the next highest level if addressed.
3. Label common errors.	Give each type of error a label that describes the error.	
4. Determine interventions.	**Step 1:** Design instruction to help learners understand the type of error and the strategies to fix that type of error. **Step 2:** Engage learners in focused revision of those types of errors in their own work.	
5. Take it to the classroom. **Note:** No student names should ever be revealed in this process.	**Strategy 1:** Swap card sets with a colleague. In your classroom, give cards to students and have students repeat the process you used to create stacks, ending with a class-generated list of types of errors made when engaging in given concepts or processes. **Strategy 2:** If the cards have student names on them, pull a few significant examples from other classrooms and use as class teaching tools, naming the errors, identifying the strategies to fix or avoid such errors, and collectively modifying or altering the examples with classroom input.	

- Erkens, 2019

Example of Common Assessments in Career & Technical Education Courses

Common Career Technical Core (12 shared standards):

https://cte.careertech.org/sites/default/files/CCTC_Standards_Formatted_2014.pdf

8. Utilize critical thinking to make sense of problems and persevere in solving them.
Career-ready individuals readily recognize problems in the workplace, understand the nature of the problem, and devise effective plans to solve the problem. They are aware of problems when they occur and take action quickly to address the problem. They thoughtfully investigate the root cause of the problem prior to introducing solutions. They carefully consider the options to solve the problem. Once a solution is agreed upon, they follow through to ensure the problem is solved, whether through their own actions or the actions of others.

Learning Target: I can solve problems. This means I can identify the appropriate concerns and the plausible causes for those concerns within a given problem and then identify possible solutions to address the problem. I can prioritize to select the right option and then gather confirmation after I apply the solution to guarantee that I have applied a workable solution.

Career & Technical Education Courses:
- Family and Consumer Science
- Agriculture
- Welding / Metals
- Construction / Woodworking
- Mechanical Engineering / Automotive
- Architectural Drafting (CAD)

** Ideas for the 4 C model are attributed to CTE teachers in West Delaware, IA*

Core Processes for Thinking Curriculum	
Problem Solving = Concern + Cause + Correction + Confirmation.	
Concern:	What is the root of the problem? Why is it a concern?
Cause:	What are the variables that created or complicate the concern?
Correction:	What are some plausible solutions, and which one is the best one to solve the problem?
Confirmation:	What evidence do you have or will you generate to ensure your solution accurately addressed the concern and solved the problem?

Erkens, In press. *The Handbook* for *Collaborative Common Assessments.*

Common Formative Assessments:

Each step of each major "C" in the problem solving model can be formatively assessed within the appropriate content area and at multiple times and places. For example, to demonstrate skills with isolating and addressing *concerns,* students must be able to do each of the following:

1. Define *concern*.
2. Determine what criteria will be used to identify something as a concern.
3. Identify possible concerns in given scenarios (discipline specific examples).
4. Identify the main concern in a discipline-specific complex problem.
5. Isolate the concern and validate that it is problematic by naming the indicators of the concern and the status of excellence that is not being met because the concern exists.

Use small exit tickets to test each C as needed within a probable (real) and currently linked curriculum concept. E.g. What does 'Concern' look like in the current foods unit? The cabinetry unit? The engines unit? Etc.

For example, the following exit tickets could be used to determine whether or not learners can identify the *appropriate causes* when solving discipline specific problems:

Task: all prompts asked learners to *name the plausible causes that you should explore so you can generate possible solutions to avoid the problem in the future.*

Family and Consumer Science
You used the attached recipe and followed the directions perfectly, but your homemade bread is not rising. Name the plausible causes.

Mechanical Engineering / Automotive
You bought a 10 year old car and it ran perfectly when you drove it home and parked it outside for the night. Today it doesn't start. Name the plausible causes.

Woodworking You measured correctly for a corner cabinet, but the cabinet you have built to match the measurements is not fitting tightly into the assigned corner. Name the plausible causes.

Agriculture
You planted a field of soybeans, using the highest caliber of seed and following the requirements perfectly. You've had ample sun and rain, but the beans are not growing. Name the plausible causes.

Erkens, 2019. *The Handbook* for *Collaborative Common Assessments.*

Courses	Total # Students	# cards right	# cards wrong	# types of errors made
Family and Consumer Science I	53	37	16	
Family and Consumer Science II	17	11	6	
Welding / Metals	19	11	8	
Construction / Woodworking	27	14	13	
Mechanical Engineering / Automotive	15	8	7	
Architectural Drafting (CAD)	21	16	5	
Agriculture	15	10	5	

Intervention Planning

Student Identifiers (e.g. A2, B1, C3, etc.)	Types of Error	Instructional Fix?
A2, A4, A10, C3	Incomplete or inaccurate list of plausible causes	
A9, B1	Cause is confused with concern.	
A3, A6, B2, B6, B7, B8, C4, C5, C10, D1	Details are confused as causes.	
A7, A8, B3, B9, C7, D6	Cause is named but reasoning is wrong	

Analyzing Types of Errors	
Mistake	**Error**
A mistake is the state or condition of being wrong because of a simple accident (misreading the directions or missing a key word like *always* or *not*). Clear additional evidence demonstrates that key concepts, terms, or processes are understood.	An error is the state or condition of being wrong because of a clear misunderstanding or application of a practice in concept, skill, reasoning, or any combination therein. Available evidence demonstrates a visible disconnect between what was taught and what was understood.

(Erkens, 2019)

Types of Errors to Consider (Chappuis, 2009)

- Is it a reading error (mistake)?
  - Phrasing of question
  - Misinterpreting directions
  - Skipping key words like *not*, *always*, *most*, and so on

- Is it a concept error (misunderstanding of the big idea or specific identifying factors)?
  - Vocabulary concerns and background knowledge
  - Which concepts? What *is* understood about the concept? What is missing?
    - Knowledge errors (unclear about definitions)
    - Unable to identify individual parts
    - Unable to explain key processes
    - Unable to link together the interworking relationships

- Is it a reasoning error (poor application of a thinking skill applied to concepts)?
  - Understanding the reasoning process
  - Employing the reasoning process inaccurately or insufficiently
    - Unsupported claims
    - Insufficient evidence and sampling errors
    - Overgeneralizations or oversimplifications
    - Inconsistency (in evidence or application)
    - Omissions
    - Contradictions
    - Illogical thinking and non-sequitur errors

Types of Errors Found in Argumentation Student Artifacts		
Making a Claim	**Employing Evidence**	**Building an Argument**
• Launches communication without first asserting anything • Creates a topic sentence rather than a claim • Creates a claim, but the stance is unclear	• Inaccurate evidence • Insufficient evidence • Summarized text • Sufficient evidence but weakly linked to the claim	*Based on your past experience, what types of errors do learners make when they struggle with organization? Write a few ideas here:*

Looking at Assessment Data

Looking for valuable information from classroom assessments involves much more than stopping at traditional scoring schemes:

- Establishing total correct out of the total possible: 12/15
- Determining passing based on pre-established cut scores like 80%
- Averaging scores and calculating based on the 100 Point Scale
- Turning rubric levels into point values rather than levels of proficiency
- Calculating totals with every question as "equal" when there are different levels of rigor involved
- Tracking point accumulation across multiple assessments

What does learning look like and how do we know it when we see it? What's sufficient? Is 12/15 (80%) a *passing* mark? Does the learner demonstrate adequate levels of proficiency based on a cut score?

Assessment Blueprint: 15 points

Test construct	**Learning Target 1**		**Learning Target 2**		**Learning Target 3**	
DOK 1: Identify	#1	MC	#6	T/F	#11	Label
DOK 1: Describe	#2	Short Answer	#7	Short Answer	#12	Short Answer
DOK 2: Compare	#3	Analogy	#8	Match	#13	Venn
DOK 2: Interpret	#4	Data Chart Labels	#9	Short Answer	#14	Data Chart Labels
DOK 3: Construct	#5	Create Model	#10	Argue with Evidence	#15	Critique

All earned 12 / 15:

*check the number right per target

*check the DOK levels of the numbers missed

Students	**3 Wrong:**			**Did this learner demonstrate proficiency? Explain your thinking:**
Student A	#13	#14	#15	
Student B	#5	#10	#15	
Student C	#2	#9	#12	

Rigor: A balance of conceptual understanding, procedural fluency and application.

Webb's Depth of Knowledge Framework (2005)

- **Level 1: Recall**– Recall of a fact, information, or procedure (Recite, Recall, Label, Naming, Define, Identify, Match, List, Draw, Calculate).
- **Level 2: Skill/Concept** - Use information or conceptual knowledge, two or more steps, etc. (Infer, Identify Patterns, Modify, Predict, Distinguish, Compare).
- **Level 3: Strategic Thinking** – Requires reasoning, developing plan or a sequence of steps, some complexity, more than one possible answer (Assess, Revise, Critique, Draw Conclusions, Differentiate, Formulate, Hypothesize, Cite Evidence).
- **Level 4: Extended Thinking** – Requires an investigation, time to think and process multiple conditions of the problem (Synthesize, Analyze, Prove, Connect, Design, Apply Concepts).

DOK requires looking at the verb in context – not the verb in isolation. The context determines the level of rigor.

Additional rigor frameworks (references in bibliography)**:**

- The New Taxonomy of Educational Objectives (Marzano & Kendall, 2007)
- Bloom's (1956) Revised Taxonomy (Anderson & Krathwohl, 2001)
- Rigor/Relevance Framework (International Center for Leadership in Education, 2012)
- Authentic Intellectual Work Framework (Newmann, King, & Carmichael, 2007)
- Hess Cognitive Rigor Matrices (Combines Bloom and Webb, by Karin Hess, 2014)

Difficulty: The difficulty of an item varies based on whether or not students found the task to be easy or hard. The level is determined using a formula based on how many got the answer to the task correct.

Rigor and Difficulty: An assessment task can have both a rigor level and a difficulty index. Rigor levels remain stationary, but difficulty levels vary. In other words, a level 3 task is level 3 for *everyone*, whereas the difficulty indices will change based on the students' readiness and overall experience with the task (some students will find the task easy while others might think it's hard). For example, imagine all students have the task of pushing the same rock up hill. The size, weight, and shape of the rock as well as the slope of the hill (rigor) is the same for everyone, but the challenge of the experience (difficulty) will vary; some people will find the task harder than others based on their own degree of physical strength and fitness.

Learning Target: I can describe the life cycle.

DOK 1	*Describe* the four stages of the Monarch butterfly's first generation lifecycle.
DOK 2	*Describe* the difference between the first and fourth generations of the Monarch butterfly's lifecycle over the course of a single year.
DOK 3	*Describe* a model you might use to represent the relationships between the lifecycle of a Monarch butterfly and its migration patterns.
DOK 4	*Describe* your research findings of the migratory patterns of the North American Monarch butterfly over the last several years. Share your data using visual tools and then use your tools to predict the long-term impact environmental shifts may likely have on the Monarch butterfly's lifecycle. Conclude your description with a plan of action to preserve the Monarch butterfly.

You Try It: Select a learning target and create questions, prompts, or tasks that involve each level of DOK for the key learning verb involved.

ELA	Math	Science	Social Studies	Music/Art	PE
Predict	Count	Defend	Analyze	Arrange	Walk
Infer	Add	Hypothesize	Generalize	Accompany	Throw
Compare/contrast	Subtract	Explain	Recognize	Create	Collaborate
Cite	Multiply	Organize	Influence	Compose	Monitor
Describe	Divide	Question	Justify	Decorate	Play
Classify	Extend	Demonstrate	Argue	Draw	Run
Explain	Calculate	Conclude	Debate	Paint	Jump

DOK level of the targets being assessed	*Verb / Learning Target:*
Question or Task:	
Level 4: Extended Thinking	
Level 3: Strategic Thinking	
Level 2: Skill or Concept	
Level 1: Recall	

Working Team Document for Team Time:
https://docs.google.com/document/d/1aimqikO1W16Nw3pH7psNic7oTPnEJjjboeZVH3J3R4k/copy

> *"We can find no evidence that any country that leads the world in educational performance has gotten there by implementing any of the major agenda items that dominate educational reform in the United States."*
>
> —Tucker, *Standing on the Shoulder of Giants: An American Agenda for Educational Reform* (2011), p. 39

"We're not teaching kids to chase their dreams; we're teaching them to catch them."

— C.J Huff, Superintendent, Joplin Missouri Public Schools

You are the ones we've been waiting for…

Hattie (2009, pp. 238–239) writes: "The conclusions [of over 800 research-based meta-analyses] are recast here as six signposts toward excellence in education:

1. Teachers are among the most powerful influences in learning.
2. Teachers need to be directive, influential, caring, and actively engaged in the passion of teaching and learning.
3. Teachers need to be aware of what each and every student is thinking and knowing, to construct meaning and meaningful experiences in light of this knowledge and have proficient knowledge and understanding of their content to provide meaningful and appropriate feedback such that each student moves progressively through the curriculum levels.
4. Teachers need to *know the learning intentions* and success criteria of their lessons, know *how well they are attaining* these criteria for all students, and know *where to go next* in light of the gap between students' current knowledge and understanding and the success criteria of: "Where are you going?", "How are you going?", and "Where to next?".
5. Teachers need to move from the single idea to multiple ideas, and to relate and then extend these ideas such that learners construct and reconstruct knowledge and ideas. It is not the knowledge or ideas, but the learner's construction of this knowledge and these ideas that is critical.
6. School leaders and teachers need to create school, staffroom, and classroom environments where error is welcomed as a learning opportunity, where discarding incorrect knowledge and understandings is welcomed, and where participants can feel safe to learn, relearn, and explore knowledge and understanding.

In these six signposts, the word 'teachers' is deliberate, as a major theme is when teachers meet to discuss, evaluate, and plan their teaching in light of the feedback evidence about the success or otherwise of their teaching strategies and conceptions about progress and appropriate challenge. This is not critical reflection, but *critical reflection in light of evidence* about their teaching."