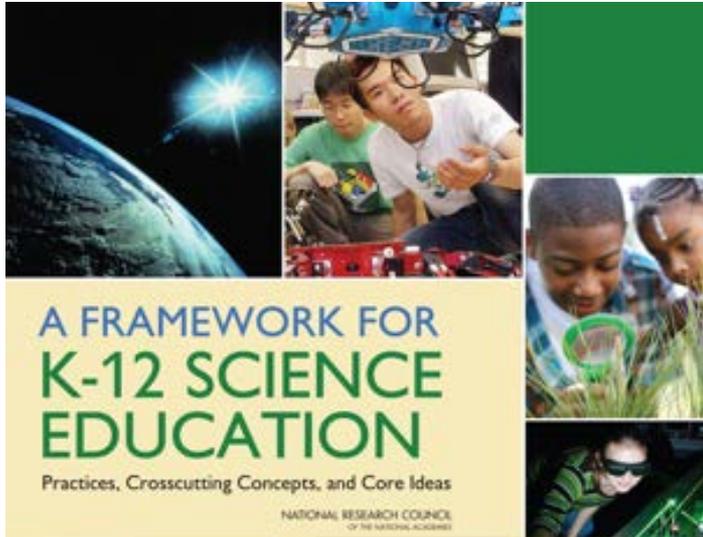


Reconceptualizing Alignment: Criteria to guide alignment and quality of NGSS assessments

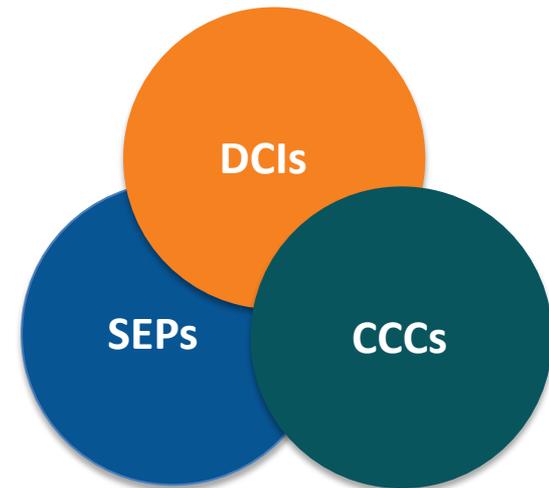


Aneesha Badrinarayan, Associate Director, Science

Framing the challenge



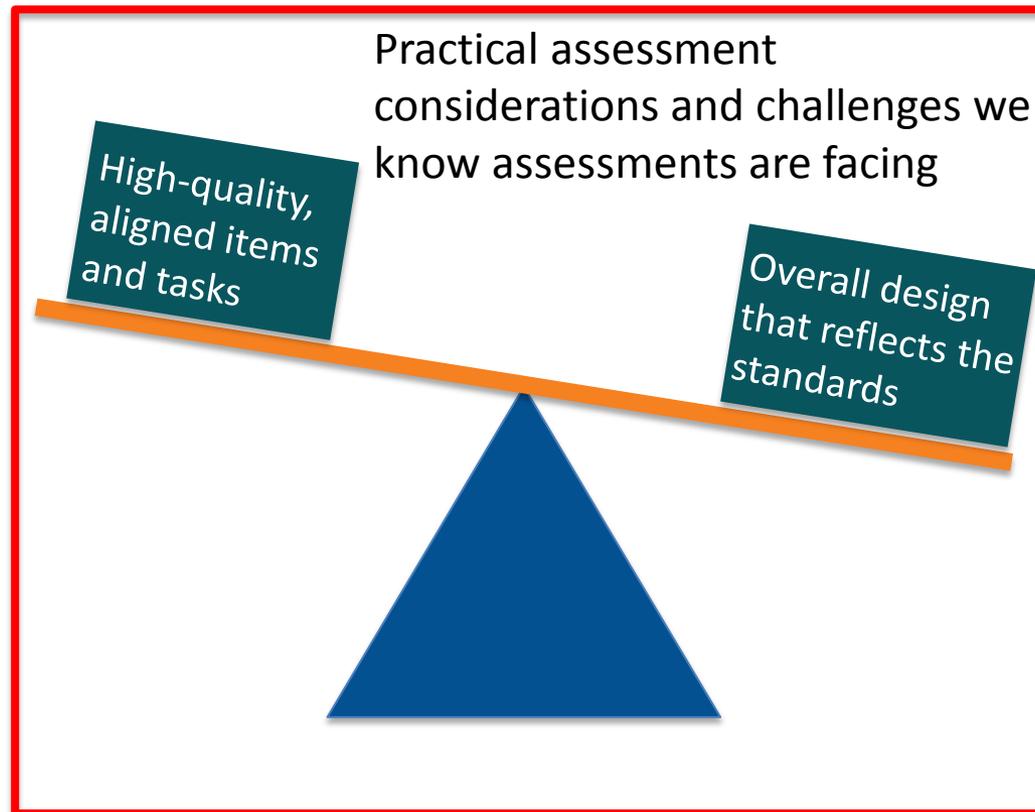
- ❖ Shifting what it means for a student to demonstrate they “know science”
 - ❖ Purposeful application of knowledge and practice
 - ❖ “knowing” shifts from recall to using scientific principles, skills, and behaviors to make sense of the world and address real-world problems
 - ❖ Application and reasoning shifts from the expectation for some standards, some students, some performances to all standards, all students, all performances



New Criteria

- With ongoing input and feedback from NGSS and Framework writers, state science and assessment leaders, and research and measurement partners, Achieve is developing criteria to help states talk about, develop, and evaluate assessments that truly reflect the intent of the standards.

Alignment Criteria Development



Overview of Criteria

1. Connection to standards through assessment design
2. Yield valuable information about student progress toward the NGSS.
3. Making sense of phenomena and designing solutions to problems
4. Reasoning with evidence grounded in the three dimensions
5. Balance across the disciplines and a range of dimensions
6. Range of cognitive complexity
7. High-quality items

Major Features

- Reflect the intent and focuses of the standards and *Framework*, including those features that can be obscured with a straight reading of individual PEs (e.g., phenomena and problems)
- Balance shared expectations for alignment with flexibility for different claims and approaches
- Contextualized by a state’s intentional assessment design, if appropriate.
- Answer the questions “what should we align to?” and “how much is enough?”, both within a given standard/target and across the range of PEs.
- Provide a concrete working definition for integration

PEs, elements, or something in between?

5-ESS1-2. Represent data in graphical displays to reveal patterns of daily changes in length and direction of shadows, day and night, and the seasonal appearance of some stars in the night sky. [Clarification Statement: Examples of patterns could include the position and motion of Earth with respect to the sun and selected stars that are visible only in particular months.] [Assessment Boundary: Assessment does not include causes of seasons.]

SEP: Analyzing and interpreting data

Represent data in graphical displays (bar graphs, pictographs and/or pie charts) to reveal patterns that indicate relationships.

DCI: Earth and the Solar System

The orbits of Earth around the sun and of the moon around Earth, together with the rotation of Earth about an axis between its North and South poles, cause observable patterns. These include day and night; daily changes in the length and direction of shadows; and different positions of the sun, moon, and stars at different times of the day, month, and year.

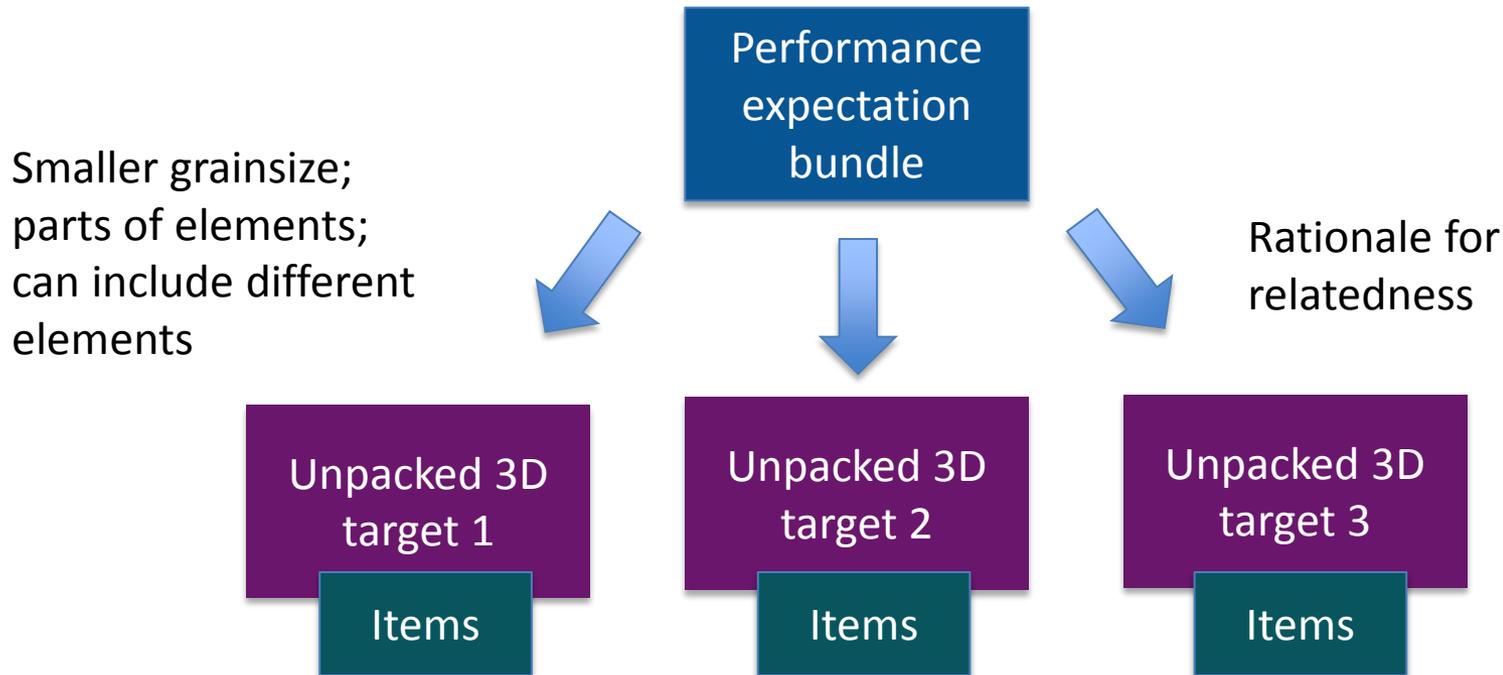
CCC: Patterns

Similarities and differences in patterns can be used to sort, classify, communicate, and analyze simple rates of change for natural phenomena

PEs, elements, or something in between?

- Fundamentally, the standards are the standards; the criteria establish the expectation that assessment items provide evidence that connect student performance to proficiency on the body of the standards
 - Maintain progressions and grade-appropriate expectations
 - Maintain consistency with adopted standards
- They reject, however, that this has to happen by asking students to demonstrate only the combinations of elements—and the PE language itself. Instead, they allow for flexibility
 - Requires careful unpacking of standards
 - Acknowledges that:
 - Relatedness among SEPs and broad usefulness of CCCs—sometimes additional SEPs and CCCs might be needed
 - Some elements might be more appropriate than others for a given test platform, context, and part of the understanding targeted
 - Some states will want to assess a more expansive conceptualization of the standards, that includes different SEP and CCC elements to be consistent with teaching and learning and goals.

PEs, elements, or something in between?



****Allows flexibility to provide valid evidence for a range of state claims and priorities while maintaining a justifiable connection back to standards**

What does it look like to assess a multidimensional target?

Unpacked 3D
target 1

Items

Defining integration

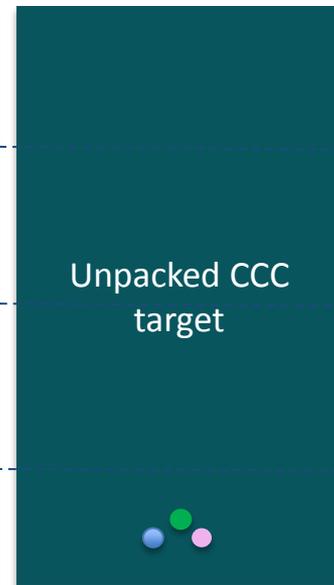
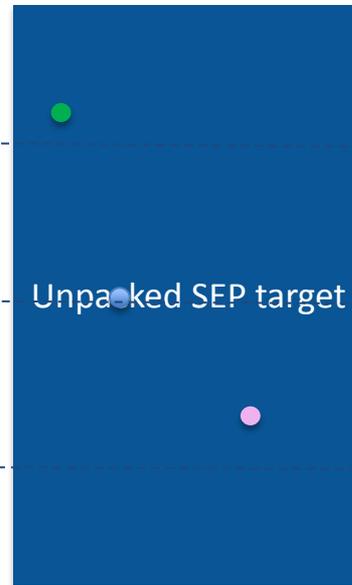
Items claiming alignment to multiple dimensions **cannot be successfully completed without using those dimensions at a grade-appropriate level of sophistication**

- distinguish between items that have a connection to a dimension and those actually assessing student knowledge and ability to use the target

Multi-dimensional items require students to **use reasoning with evidence and scientific principles to address a phenomenon or problem.**

What does it look like to assess a multidimensional target?

Fully requires the targeted knowledge and practice (i.e., cannot complete the task without complete understanding)



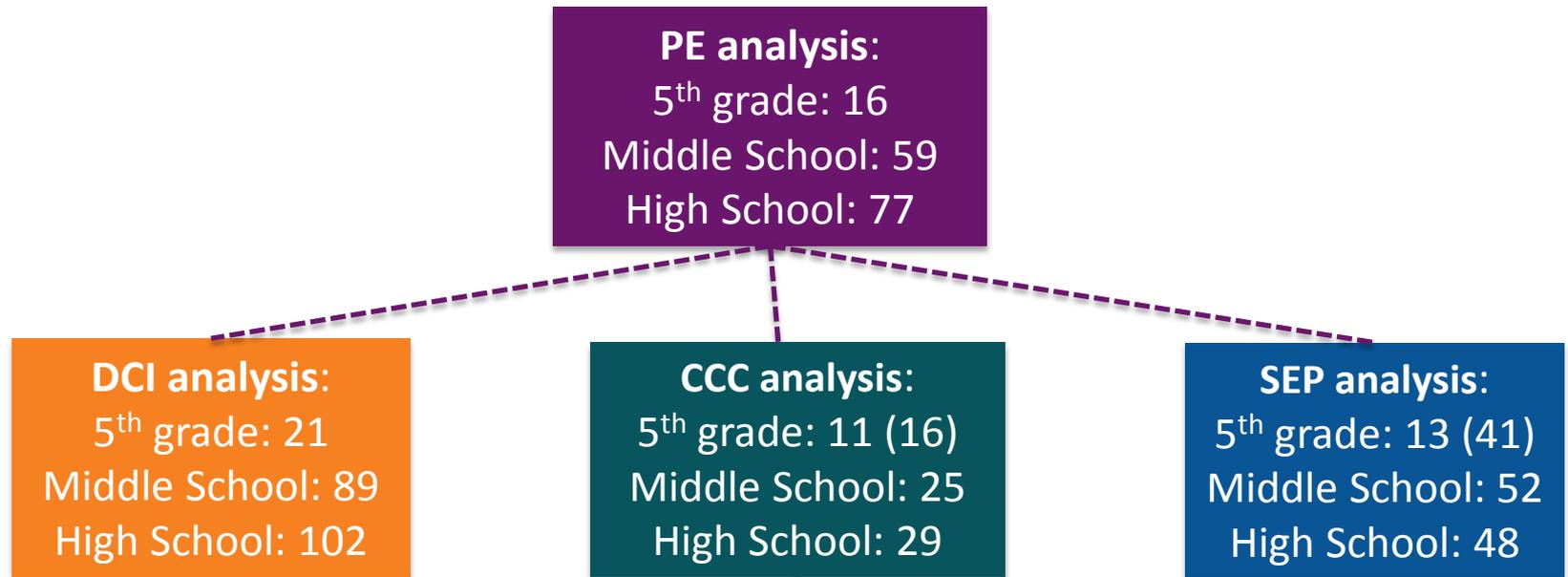
Does not require the targeted knowledge and practice at all (i.e., no understanding is needed to complete the task/portion of the task; a student could complete the task without any knowledge of the three dimensions)

Integrating the three dimensions by focusing on reasoning with evidence and scientific principles

- Differentiator between assessment items that embodied the intent of a given PE or other 3D target: reasoning with evidence and scientific principles.
- Four categories of reasoning that can be appropriate:
 - Reasoning to interpret information to generate evidence [application of DCIs, SEPS, and CCCs]
 - Reasoning with provided evidence [foregrounds application of SEPs with opportunity to engage DCIs and CCCs]
 - Reasoning with scientific principles [foregrounds DCIs and CCCs while engaging SEPs]
 - Reasoning with scientific evidence and principles [emphasizes SEPs, CCCs, DCIs]
- Contextualized by the expectations about alignment to standards we discussed previously, this criterion provides a more concrete target for assessment items and tasks.
- A note about engaging CCCs

How much is enough?

I. Sampling



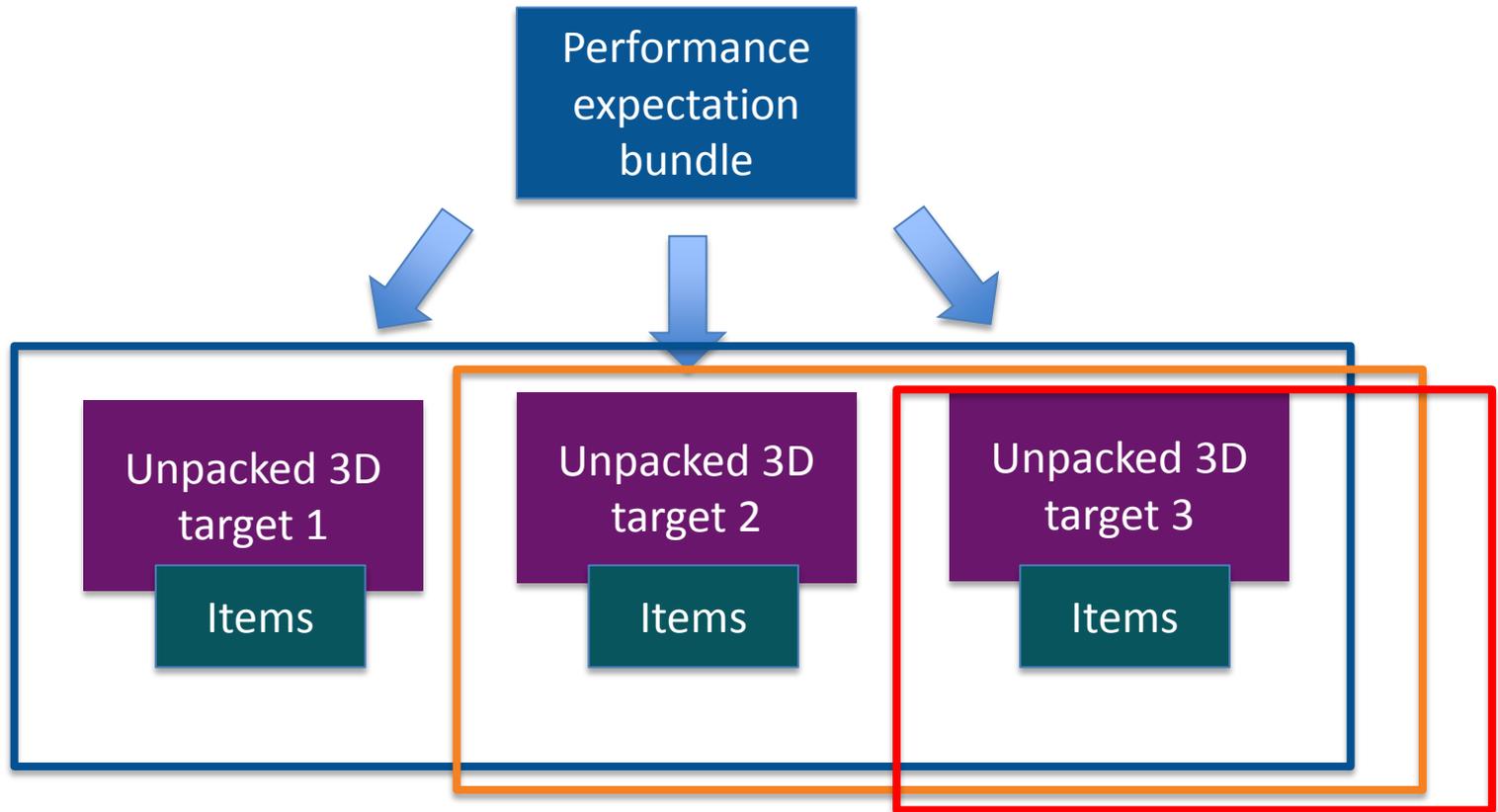
If you can't assess it all, what needs to be there?

How much is enough?

- Roughly mirror the disciplinary distribution of the targeted grade level or band of standards, unless there is an intentional, explicit design decision that would skew this (e.g., a series of EOC assessments)
 - Allows flexibility with regard to DCI distribution, but ensures that all disciplines are valued
- Include multiple elements of at least 4 SEPs and 4 CCCs
 - remember, this is the baseline
 - Intended to ensure a variety of SEPs and CCCs are assessed, but provide flexibility to meet a particular state's claims and reporting needs

How much is enough?

II. Within a standard



- Does not need to be comprehensive
- Does need to be sufficient and focus on the parts of the PEs with the most explanatory (and revelatory) value
 - Individual items should assess core components of the targeted understanding
 - Collectively, all of the items with a common target should provide sufficient evidence of student proficiency
 - Relies on consensus amongst experts

Next steps

- Criteria are designed to provide a common understanding and language for us, as a community, to use to build the best possible assessments from a content perspective.
- Full criteria, including evidence descriptors, sample evaluation guidance, examples, and tradeoffs to consider are undergoing review
- Our focus to this point has been content; we'd like to continue this focus and expand to other considerations
- If you are interested in being part of the review process, or part of a focus group to discuss the criteria more deeply, please let me know!

Contact information

Aneesha Badrinarayan
Associate Director, Science
abadrinarayan@achieve.org

Questions?