Using a Novice-Expert Perspective to Structure State and Classroom Assessment Claims/Interpretations of the NGSS

Brian Gong
Center for Assessment
Mary Norris
Virginia Tech

Presentation at the
NCME Classroom Assessment Conference
October 9, 2018 Lawrence, KS

Center for Assessment

My Central Messages

- We need to create more specific ***claims*** regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.
- The claims may be thought of as assertions about the development and application of ***expertise***, along several dimensions.
  - The NGSS in particular require taking a stance about *transfer* and *generalization* because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.
- Researchers and developers have come up with ***examples*** of how science expertise can be further defined and assessed
- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

My Central Messages

- We need to create more specific ***claims*** regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.

- The claims may be thought of as assertions about the development and application of ***expertise***, along several dimensions.
  - The NGSS in particular require taking a stance about *transfer* and *generalization* because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.

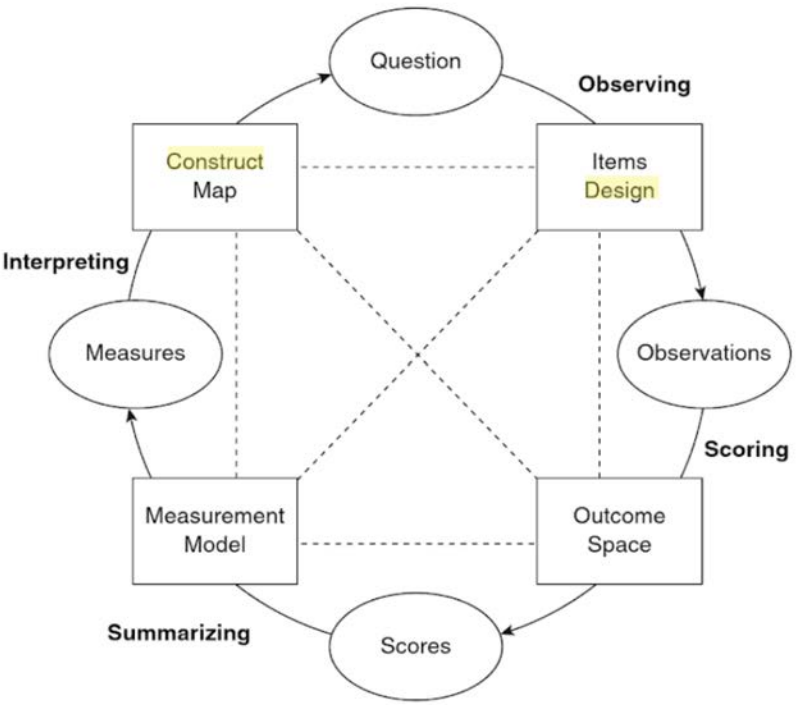
- Researchers and developers have come up with ***examples*** of how science expertise can be further defined and assessed

- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

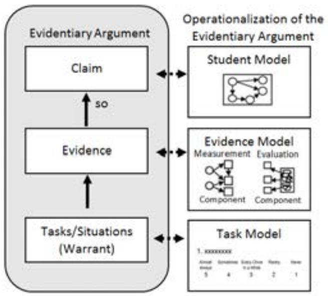
The centrality of claims in assessment

- A ***claim*** is an assertion, an ***intended interpretation*** of student ability or performance, with intended ***uses***
  - "The student can/has…"
- The claim guides the ***design and construction*** of the assessment
  - What *evidence* would be needed to support/falsify the claim?
  - "Evidence-centered design" (Mislevy et al.), "construct modeling" (Wilson et al.)
- Intended interpretations are the focus of ***validation*** efforts in assessment
  - "Interpretive Argument"; "Validation Argument" (Kane)
    - Supported interpretation in movement from observation to claim about non-visible construct (e.g., ability) and more distal/generalized contexts/domains

Claims & modern assessment design



Source: Brown, Nagashima, Fu, Timms, & Wilson, 2010



Adapted from Mislevy & Haertel, 2006.

Table 2. Design Pattern Attributes and Corresponding Assessment Argument Components

Attribute	Value(s)	Assessment Argument Component
Rationale	Explain why this item is an important aspect of scientific inquiry.	Warrant (underlying)
Focal Knowledge, Skills, and Abilities	The primary knowledge/skill/abilities targeted by this design pattern.	Student Model
Additional Knowledge, Skills, and Abilities	Other knowledge/skills/abilities that may be required by this design pattern.	Student Model
Potential observations	Some possible things one could see students doing that would give evidence about the (knowledge/skills/attributes) KSAs.	Evidence Model
Potential work products	Modes, like a written product or a spoken answer, in which students might produce evidence about KSAs.	Task Model
Characteristic features	Aspects of assessment situations that are likely to evoke the desired evidence.	Task Model
Variable features	Aspects of assessment situations that can be varied in order to shift difficulty or focus.	Task Model

Source: Mislevy & Riconscente, 2005

Center for Assessment

Claims typically not detailed enough

- Typical sources of claims and intended interpretations in large-scale assessments
  - **Content standards** (e.g., NGSS)
  - Assessment reported scores (e.g., Science scale 200-900)
  - Assessment Performance/Achievement Level Descriptors [**PLDs/ALDs**] (e.g., “Below Basic/ Basic/Proficient/Advanced”)
  - Other supportive interpretive materials (e.g., example items and annotated student work, test blueprints, learning progressions, curricula, predictive relationships to other valued performances/constructs)

My Central Messages

- We need to create more specific ***claims*** regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.

- The claims may be thought of as assertions about the development and application of ***expertise***, along several dimensions.
  - The NGSS in particular require taking a stance about *transfer* and *generalization* because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.

- Researchers and developers have come up with ***examples*** of how science expertise can be further defined and assessed

- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

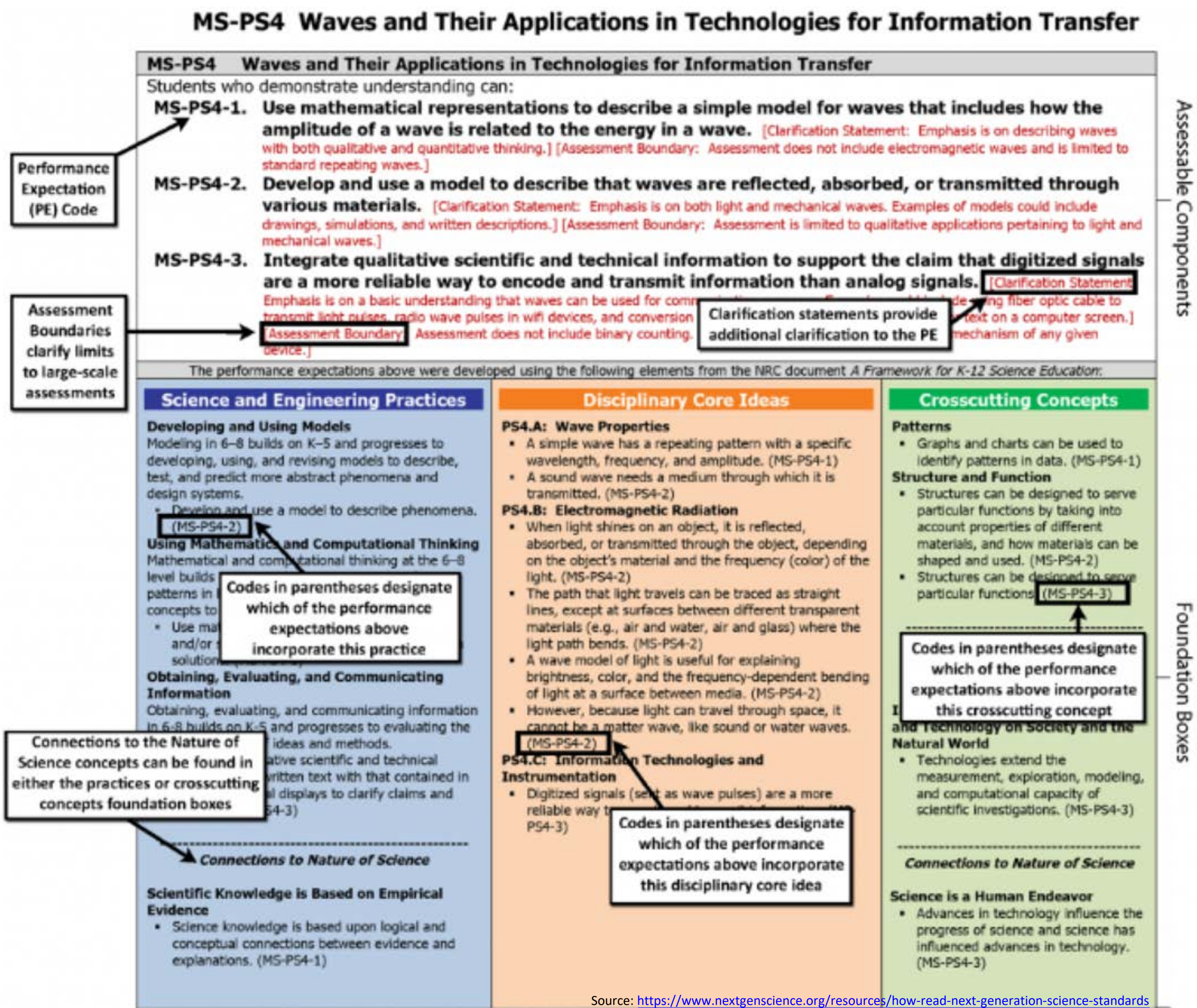
Example: Possible growth/expertise claims

Dimension	Description of type of growth
Independence	“The student was able to do similar tasks with less support”
Correctness	“The student mastered things s/he had incorrect previously”
Curricular sequence	“The student learned things in the next curricular sequence”
Near transfer (similar examples)	“The student addressed new, similar examples of the same knowledge and skills”
Cognitive complexity of same knowledge/skills	“The student addressed more complex examples of the same knowledge and skills”
Fluidity & confidence	“The student became quicker and more confident in applying the same knowledge and skills”
Far transfer (application) of knowledge and skills	“The student can apply knowledge and skills to new, different situations, generating new combinations and solutions”
Expert knowledge representations & skills	“The student learned more sophisticated and refined knowledge, and applied skills more expertly”
Internal, creative roles, purposes, & values	“The student shifts roles to generate her/his own problems, purposes, according to own values consistent with discipline”

Example: Possible growth/expertise claims - 2

Dimension	Description of type of growth
Ease of learning	“The student is able to learn new related knowledge and skills”
Memory	“The student is able to perform long after initially learning”
Reflection/monitoring	“The student is able to self-correct own understanding & actions”
Connection	“The student is able to make fruitful connections between parts”
Persistence	“The student persists through challenges”
Creativity	“The student is able to invent/apply different approaches”
Teach	“The student is able to communicate clearly to help others”
Open	“The student considers others’ information and values”
Collaboration	“The student works with others to accomplish shared goals in ways reflective of discipline/real world”

Sample NGSS Performance Expectation



MS-PS4 Waves and Their Applications in Technologies for Information Transfer

MS-PS4 Waves and Their Applications in Technologies for Information Transfer

Students who demonstrate understanding can:

MS-PS4-1. Use mathematical representations to describe a simple model for waves that includes how the amplitude of a wave is related to the energy in a wave. [Clarification Statement: Emphasis is on describing waves with both qualitative and quantitative thinking.] [Assessment Boundary: Assessment does not include electromagnetic waves and is limited to standard repeating waves.]

MS-PS4-2. Develop and use a model to describe that waves are reflected, absorbed, or transmitted through various materials. [Clarification Statement: Emphasis is on both light and mechanical waves. Examples of models could include drawings, simulations, and written descriptions.] [Assessment Boundary: Assessment is limited to qualitative applications pertaining to light and mechanical waves.]

MS-PS4-3. Integrate qualitative scientific and technical information to support the claim that digitized signals are a more reliable way to encode and transmit information than analog signals. [Clarification Statement: Emphasis is on a basic understanding that waves can be used for comm... ...ing fiber optic cable to transmit light pulses, radio wave pulses in wifi devices, and conversion ... text on a computer screen.] [Assessment Boundary: Assessment does not include binary counting. ... mechanism of any given device.]

The performance expectations above were developed using the following elements from the NRC document *A Framework for K-12 Science Education*:

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Developing and Using Models Modeling in 6–8 builds on K–5 and progresses to developing, using, and revising models to describe, test, and predict more abstract phenomena and design systems. • Develop and use a model to describe phenomena. (MS-PS4-2) **Using Mathematics and Computational Thinking** Mathematical and computational thinking at the 6–8 level builds ... patterns in ... concepts to ... • Use ma... and/or ... solution... **Obtaining, Evaluating, and Communicating Information** Obtaining, evaluating, and communicating information in 6-8 builds on K-5 and progresses to evaluating the ... ideas and methods. • ...ative scientific and technical ...ritten text with that contained in ...l displays to clarify claims and ...4-3) *Connections to Nature of Science* **Scientific Knowledge is Based on Empirical Evidence** • Science knowledge is based upon logical and conceptual connections between evidence and explanations. (MS-PS4-1)	**PS4.A: Wave Properties** • A simple wave has a repeating pattern with a specific wavelength, frequency, and amplitude. (MS-PS4-1) • A sound wave needs a medium through which it is transmitted. (MS-PS4-2) **PS4.B: Electromagnetic Radiation** • When light shines on an object, it is reflected, absorbed, or transmitted through the object, depending on the object's material and the frequency (color) of the light. (MS-PS4-2) • The path that light travels can be traced as straight lines, except at surfaces between different transparent materials (e.g., air and water, air and glass) where the light path bends. (MS-PS4-2) • A wave model of light is useful for explaining brightness, color, and the frequency-dependent bending of light at a surface between media. (MS-PS4-2) • However, because light can travel through space, it cannot be a matter wave, like sound or water waves. (MS-PS4-2) **PS4.C: Information Technologies and Instrumentation** • Digitized signals (sent as wave pulses) are a more reliable way t... (MS-PS4-3)	**Patterns** • Graphs and charts can be used to identify patterns in data. (MS-PS4-1) **Structure and Function** • Structures can be designed to serve particular functions by taking into account properties of different materials, and how materials can be shaped and used. (MS-PS4-2) • Structures can be designed to serve particular functions (MS-PS4-3) ... and Technology on Society and the Natural World • Technologies extend the measurement, exploration, modeling, and computational capacity of scientific investigations. (MS-PS4-3) *Connections to Nature of Science* **Science is a Human Endeavor** • Advances in technology influence the progress of science and science has influenced advances in technology. (MS-PS4-3)

Source: https://www.nextgenscience.org/resources/how-read-next-generation-science-standards

Specify claim

Dimension	
Independence	Ease of learning
Correctness	Memory
Curricular sequence	Reflection/monitoring
Near transfer (similar examples)	Connection
Cognitive complexity of same knowledge/skills	Persistence
Fluidity & confidence	Creativity
Far transfer (application) of knowledge and skills	Teach
Expert knowledge representations & skills	Open
Internal, creative roles, purposes, & values	Collaboration

NGSS Sampling & Balance of Emphasis

- To provide evidence to inform the Claims
  - What should be sampled? What should be the balance of emphasis?
- NGSS 3D = 39 DCI x 8 SEP x 7 CCC = **2,184 possible** unique combinations of 1 DCI/SEP/CCC: Too many to learn or assess directly (?)
- NGSS Solution: **194 Performance Expectations**, total across all the grades—very thin/focused and narrow pieces of DCI/SEP/CCC

NGSS PE Sampling

NGSS Disciplinary Content Ideas (DCI) and Scientific &Engineering Practices (SEP) Paired in the Performance Expectations, Grades 3-5

Grade	AQDP	DUM	PCOI	AID	UMCT	CEDS	EAE	OECI
Grade 3								
PS	**X**		**X**					
LS		**X**		**X**		**X**	**X**	
ESS				**X**			**X**	**X**
Grade 4								
PS	**X**	**X**	**X**			**X**		
LS		**X**					**X**	
ESS			**X**	**X**		**X**		**X**
Grade 5								
PS		**X**	**X**		X		**X**	
LS		**X**					**X**	
ESS		**X**		**X**	X		**X**	**X**
3-5 ETS	**X**		X			X		

AQDP	Asking Questions and Defining Problems
DUM	Developing and Using Models
PCOI	Planning and Carrying Out Investigations
AID	Analyzing and Interpreting Data
UMCT	Using Mathematics and Computational Thinking
CEDS	Constructing Explanations and Designing Solutions
EAE	Engaging in Argument from Evidence
OECI	Obtaining, Evaluating, and Communicating Information

NGSS PE Sampling and BoE - 2

- NGSS PEs (Component Areas x SEP x CCC), Gr. 5

Performance Expectations by DCI by SEP

	Grade 5		AQDP	DUM	PCOI	AID	UMCT	CEDS	EAE	OECI
Physical Sciences	PS1, Matter and Its Interactions	PS1.A, Structure and Properties of Matter		5-PS1-1	5-PS1-3		5-PS1-2			
		PS1.B, Chemical Reactions			5-PS1-4		5-PS1-2			
		PS1.C Nuclear Processes								
	PS2, Motion and Stability: Forces and Motion	PS2.A, Forces and Motion								
		PS2.B, Types of Interactions							5-PS2-1	
		PS2.C, Stability and Instability in Physical Systems								
	PS 3, Energy	PS3.A, Definitions of Energy								
		PS3.B, Conservation of Energy and Energy Transfer								
		PS3..C, Energy and Chemical Processes in Everyday Life								
		PS3.D Waves and Their Applications in Technologies for Information Transfer		5-PS3-1						
	PS4, Wave Properties	PS4.A, Wave Properties								
		PS4.B, Electromagnetic Radiation								
		PS4.C, Information Technologies and Instrumentation								

Performance Expectations by DCI by SEP

Grade 5		AQDP	DUM	PCOI	AID	UMCT	CEDS	EAE	OECI
, Matter and Its Interactions	PS1.A, Structure and Properties of Matter		5-PS1-1	5-PS1-3		5-PS1-2			
	PS1.B, Chemical Reactions			5-PS1-4		5-PS1-2			
	PS1.C Nuclear Processes								
lotion and Stability: rces and Motion	PS2.A, Forces and Motion								
	PS2.B, Types of Interactions							5-PS2-1	
	PS2.C, Stability and Instability in Physical Systems								
PS 3, Energy	PS3.A, Definitions of Energy								
	PS3.B, Conservation of Energy and Energy Transfer								
	PS3..C, Energy and Chemical Processes in Everyday Life								
	PS3.D Waves and Their Applications in Technologies for Information Transfer		5-PS3-1						
Wave Properties	PS4.A, Wave Properties								
	PS4.B, Electromagnetic Radiation								
	PS4.C, Information Technologies and Instrumentation								
From Molecules to sms: Structures and Processes	LS1.A, Structure and Function								
	LS1.B, Growth and Development of Organisms								
	LS1.C, Organization for Matter and Energy Flow in Organisms							5-LS1-1	
	LS1.D, Social Interactions and Group Behavior								
2, Ecosystems: ctions, Energy, and Dynamics	LS2.A, Interdependent Relationships in Ecosystems		5-LS2-1						
	LS2.B, Cycles of Matter and Energy Transfer in Ecosystems		5-LS2-1						
	LS2.C, Ecosystem Dynamics, Functioning, and Resilience								
	LS2.D, Social Interactions and Group Behavior								
LS3, Heredity: Inheritance and Variation of Traits	LS3.A, Inheritance of Traits								
	LS3.B, Variation of Traits								
LS4, Biological Evolution: Unity and Diversity	LS4.A, Evidence of Common Ancestry								
	LS4.B, Natural Selection								
	LS4.C, Adaptation								
	LS4.D, Biodiversity and Humans								

PE coverage of SEP x DCI, grade 5

Challenge 2

- When the full DCI, SEP, and CCC are listed, the sampling by PE is clearly extremely sparse

Performance Expectations by DCI by SEP

	Grade 5		AQDP	DUM	PCOI	AID	UMCT	CEDS	EAE	OECI
Physical Sciences	PS1, Matter and Its Interactions	PS1.A, Structure and Properties of Matter		5-PS1-1	5-PS1-3		5-PS1-2			
		PS1.B, Chemical Reactions			5-PS1-4		5-PS1-2			
		PS1.C Nuclear Processes								
	PS2, Motion and Stability: Forces and Motion	PS2.A, Forces and Motion								
		PS2.B, Types of Interactions							5-PS2-1	
		PS2.C, Stability and Instability in Physical Systems								
	PS 3, Energy	PS3.A, Definitions of Energy								
		PS3.B, Conservation of Energy and Energy Transfer								
		PS3..C, Energy and Chemical Processes in Everyday Life								
		PS3.D Waves and Their Applications in Technologies for Information Transfer		5-PS3-1						
	PS4, Wave Properties	PS4.A, Wave Properties								
		PS4.B, Electromagnetic								
		PS4.C, Information Technologies and Instrumentation								
Life Sciences	LS1, From Molecules to Organisms: Structures and Processes	LS1.A, Structure and Function								
		LS1.B, Growth and Development of Organisms								
		LS1.C, Organization for Matter and Energy Flow in Organisms							5-LS1-1	
		LS1.D, Social Interactions and Group Behavior								
	LS2, Ecosystems: Interactions, Energy, and Dynamics	LS2.A, Interdependent Relationships in Ecosystems		5-LS2-1						
		LS2.B, Cycles of Matter and Energy Transfer in Ecosystems		5-LS2-1						
		LS2.C, Ecosystem Dynamics, Functioning, and Resilience								
		LS2.D, Social Interactions and Group Behavior								
	LS3, Heredity: Inheritance and Variation of Traits	LS3.A, Inheritance of Traits								
		LS3.B, Variation of Traits								
	LS4, Biological Evolution: Unity and Diversity	LS4.A, Evidence of Common Ancestry								
		LS4.B, Natural Selection								
		LS4.C, Adaptation								
		LS4.D, Biodiversity and Humans								
Earth and Space Sciences	ESS1, Earth's Place in the Universe	ESS1.A, The Universe and Its Stars							5-ESS1-1	
		ESS1.B, Earth and the Solar System				5-ESS1-2				
		ESS1.C, The History of Planet Earth								
	ESS2, Earth's Systems	ESS2.A, Earth Materials and Systems		5-ESS2-1						
		ESS2.B, Plate Tectonics and Large-Scale System Interactions								
		ESS2.C, The Roles of Water in Earth's Surface Processes					5-ESS2-2			
		ESS2.D, Weather and Climate								
		ESS2.E, Biogeology								
	ESS3, Earth and Human Activities	ESS3.A, Natural Resources								
		ESS3.B, Natural Hazards								
		ESS3.C, Human Impacts on Earth Systems								5-ESS3-1
		ESS3.D, Global Climate Change								
3-5 Engineering Design	ETS1, Engineering Design	ETS1-A, Defining and Delimiting Engineering Problems	3-5-ETS1-1							
		ETS1-B, Developing Possible Solutions			3-5-ETS1-3			3-5-ETS1-2		
		ETS-C, Optimizing the Design Solution			3-5-ETS1-3					

Expertise needed: generalization, transfer

- Can overcome sparseness of PEs by ***instructing*** more broadly, and ***assessing*** narrowly
  - Assessment is a (random/purposeful?) sample that we can generalize to a larger, learned universe ("If we assessed something else, the student would have done as well…")
  - PEs are then *assessment specifications,* and not *instructional* specifications
- Or, can in the assessment require students to ***transfer*** and ***generalize*** knowledge/skills
  - Supports the claim that student could (learn) to do all, even if instructed not on all
  - SEP and CCC may be more important than DCI

My Central Messages

- We need to create more specific ***claims*** regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.
- The claims may be thought of as assertions about the development and application of ***expertise***, along several dimensions.
  - The NGSS in particular require taking a stance about *transfer* and *generalization* because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.
- Researchers and developers have come up with ***examples*** of how science expertise can be further defined and assessed
- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

Some relevant work on expertise in science assessments from the past

- Drawing on work by Shavelson, Baxter, Glaser, Wilson, Mislevy, and colleagues
  - Context was primarily science performance tasks, 1990s, many state programs and university-based R&D projects
    - Check for correspondence on "science standards" to NGSS
  - Analytical approach was primarily cognitive psychologists, most with considerable measurement expertise
  - Focus is always on student making a claim and supporting with evidence in scientific ways

Task types: Shavelson et al.'s four types (1997)

Comparative investigation

- Paper Towels: Discover which of three kinds of paper towels holds the most water and which holds the least (Baxter, Shavelson, Goldman, & Pine, 1992).
- Bubbles: Discover which of three soapy solutions produces the most durable bubbles (Solano-Flores, 1994; Solano-Flores & Shavelson, 1994b).
- Incline Planes: Determine the relationship between the angle of inclination and the amount of force needed to move an object up a plane (Solano-Flores, Jovanovic, Shavelson, & Bachman, 1994).

Component identification

- Electric Mysteries: Determine the components of the mystery box (Shavelson, Baxter, & Pine, 1991 1.
- Mystery Powders: Given a bar containing substances commonly found in the kitchen (e.g., baking soda, starch, sugar), determine which substances are in the bag (Baxter, Elder, & Glaser, 1995; Baxter & Shavelson,1995).
- Motor: Given a motor, a battery, and a box containing a battery, determine the polarity of the battery that is inside the box (Druker, Solano-Flores, Brown, & Shavelson, 1996).

Classification

- Sink & Float: Create a classification system that allows you to predict whether an object will sink or float in tap water (Solano-Flores, Shavelson, Ruiz-Primo, Schultz, Wiley, & Brown, 1997).
- Rocks & Charts: Given a set of minerals, test the minerals for known attributes and create a classification system using those attributes (Druker, 1997).

Observation

- Daytime Astronomy: Model the path of the sun from sunrise to sunset and use direction, length, and angle of shadows to solve location problems (Solano-Flores, Shavelson, Ruiz-Primo, Schultz, Wiley, & Brown, 1997).

Types of tasks & scoring systems (Shavelson et al.)

Types of science tasks	Types of scoring systems
Comparative	Student conducts an experiment to compare two or more objects on some property. The scoring system is **procedure-based**-it focuses on the scientific defensibility of the procedures used by the student to compare the objects. For example, in *Paper Towels,* the student conducts an experiment to find out which of three kinds of paper towels holds the most water and which holds the least water. If the student does not completely saturate one of the towels, even though he or she gets the right answer, the investigation is flawed.
Component identification	Student tests objects to determine their components or how those components are organized. The scoring system is **evidence-based**-it focuses on the quality of the evidence used to confirm or disconfirm the presence of components. For example, in *Electric Mysteries,* the student has to test 6 mystery boxes to determine their contents-two batteries, a wire, a bulb, a battery and a bulb, or nothing (two boxes have the same contents). A student who tests a mystery box first with a simple circuit containing a light bulb and, then, if the bulb doesn't light, tests the circuit with a battery and a bulb, uses a scientifically defensible way of confirming or dis-confirming the presence of components.
Classification	Student classifies objects according to critical attributes to serve a practical or conceptual purpose. The scoring system is **dimension-based**-it focuses on how well the classification system constructed uses attributes that are relevant to the purposes of classification. For example, in Sink and Float, the student has to construct a classification scheme based on variables (dimensions) critical to floatation and use a classification scheme to predict if a set of bottles of different volumes and masses will sink or float. To classify objects as "floaters" and "sinkers," a student should consider mass, volume, and the interaction of mass and volume.
Observation	Student performs observations and/or models a process that cannot be manipulated. The scoring system is **accuracy-based**-it focuses on the accuracy of the observations performed and the models constructed. For example, in *Daytime Astronomy,* the student has to solve location problems by modeling sun shadows and to describe what shadows look like in different locations. A correct solution to the location problems is obtained when, among other things, the student models the sunlight and the earth's rotation, respectively, by shining the flashlight on the equator and rotating the earth globe to the East.

Classification Task? – Learning Progression for "Sinking & Floating" (Wilson, BEAR)

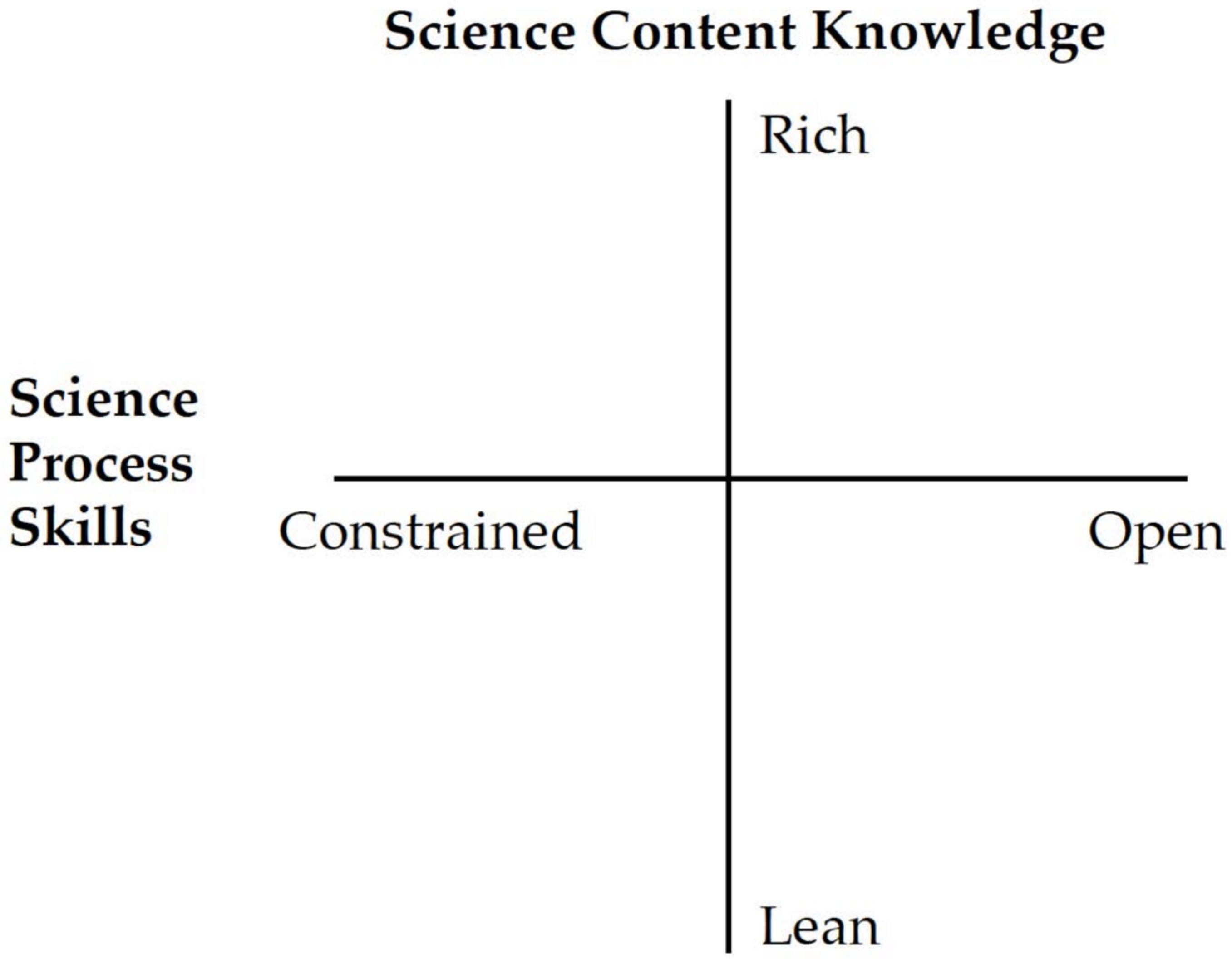
Level	What the Student Already Knows	What the Student Needs to Learn
RD	**Relative Density** Student knows that floating depends on having less density than the medium, or at least that floating depends on relative density in some way. Mentions the densities of the object and the medium.	
D	**Density** Student knows that floating depends on having less density, or at least that floating is related to density in some way.	To progress to the next level, student needs to recognize that the medium plays an equally important role in determining if an object will sink or float.
MV	**Mass and Volume** Student knows that floating depends on having less mass and more volume, or at least knows that mass and volume work together to affect floating and sinking.	To progress to the next level, student needs to understand the concept of density as a way of combining mass and volume into a single property.
M / V	**Mass** Student knows that floating depends on having less mass. **Volume** Student knows that floating depends on having more volume.	To progress to the next level, student needs to recognize that changing EITHER mass OR volume will affect whether an object sinks or floats.
UF	**Unconventional Feature** Student thinks that floating depends on an unconventional feature, such as shape, surface area, or hollowness.	To progress to the next level, student needs to rethink their ideas in terms of mass and/or volume. For example, hollow objects have a lot of volume but not a lot of mass.
OT	**Off Target** Student does not attend to any property or feature to explain floating.	To progress to the next level, student needs to focus on some property or feature of the object in order to explain why it sinks or floats.
NR	**No Response** Student left the response blank.	To progress to the next level, student needs to respond to the question.

Comparative investigation task template for varying independence (Shavelson et al.)

Shell for developing Comparative Investigations Tasks, "Low/High Inquiry levels"			
Less independence on inquiry		**More independence on inquiry**	
Step 1	Provide preparatory knowledge in one of three ways: * Written instruction * Illustration with related task *Illustration with embedded task	Step 1	Introduce the concepts that will be used in the assessment.
Step 2	Pose a problem or a hypothesis involving one relevant independent variable .	Step 2	Pose a problem or a hypothesis involving one relevant independent variable (A) and one irrelevant independent variable (B).
Step 3	Provide equipment-include independent variable. Introduce variable name.	Step 3	Provide equipment-include independent variable A and independent variable B. Introduce variable names.
Step 4	Tell the students which manipulations should be done and how they should be done.	Step 4	Ask the students to solve the problem or test the hypothesis.
Step 5	Ask students to solve the problem or test the hypothesis.	Step 5	Ask students to report manipulations, measurements, and results.
Step 6	Ask students to report manipulations, measurements, and results. Provide table/chart.		

Center for Assessment

Science Content x Process complexity matrix (Glaser, 1997)



Science Content x Process complexity – 2
(Glaser, 1997)

	Content Lean	Content Rich
Process Open	E.g., require students to coordinate a sequence of process skills with minimal demands for content knowledge. Students structure the problem in terms of actions that follow from what they know about the [specific task]. They then implement a strategy, and revise their strategy, if necessary, based on task feedback.	E.g., identification of the causal variables involved requires substantial knowledge of physics concepts of force and motion, the ability to design and carry out controlled experimentation, and the effective employment of model-based reasoning skills
Process Constrained	E.g., require minimal prior knowledge or school experiences with subject specific concepts and procedures to successfully complete the task. Rather, students are guided to carry out a set of procedures and then asked to respond to a set of questions about the results of their activities.	E.g., emphasize knowledge generation or recall—that is, "knowing" science versus "doing" science. [For example,] a comprehensive, coherent explanation revolves around a discussion of inputs, processes, and products

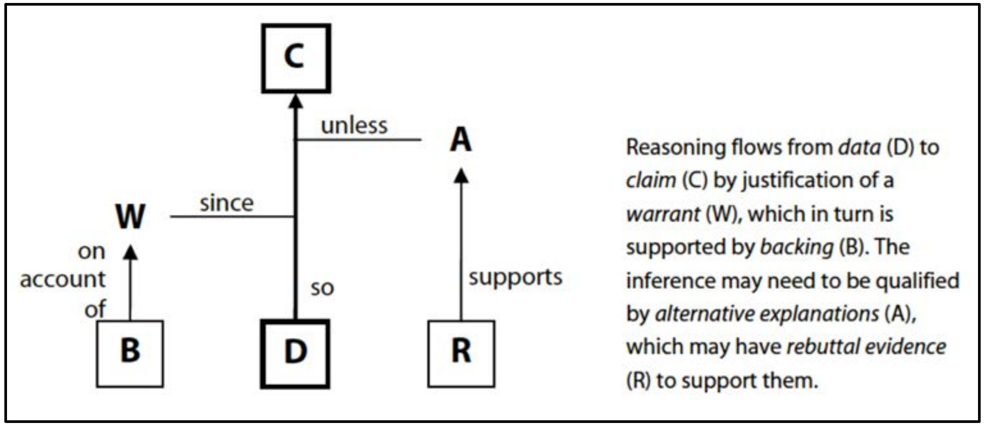
Science Content x Process complexity

(Songer)

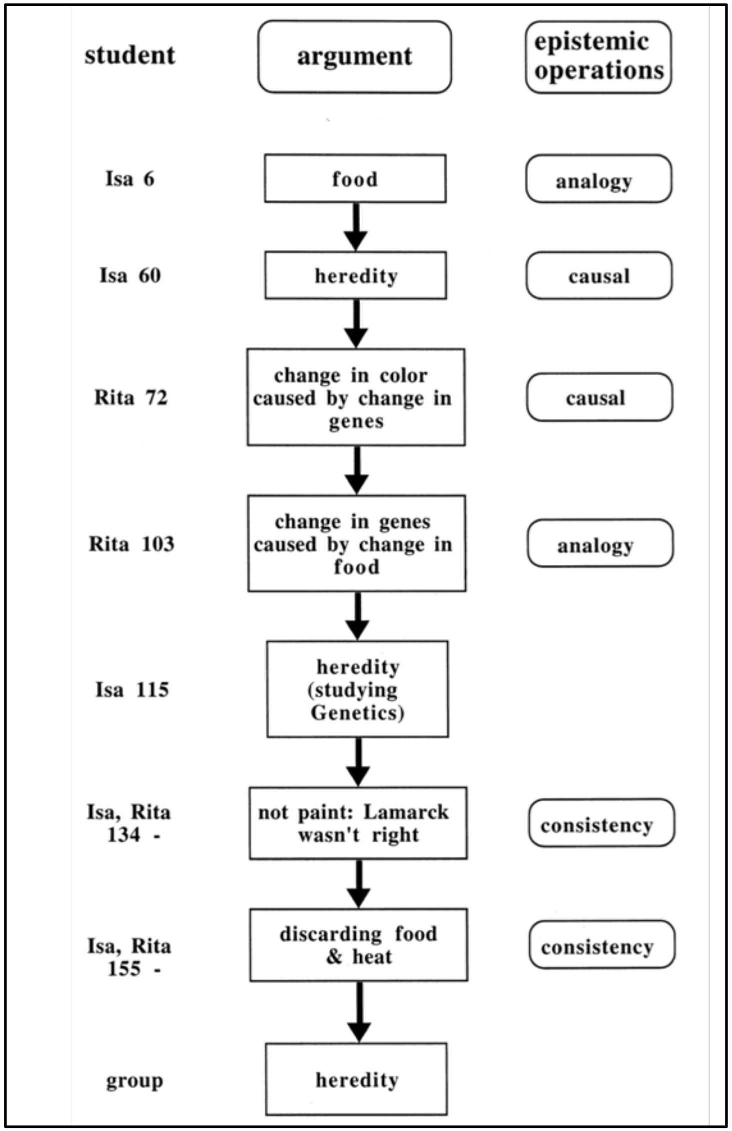
Levels of content and inquiry for tasks focused on “formulating scientific explanations from evidence”

	Content Complexity		
Inquiry Level	**Simple** – minimal or no extra content knowledge is required and evidence does not require interpretation	**Moderate** - students must either interpret evidence or apply additional (not given) content knowledge	**Complex** – students must apply extra content knowledge and interpret evidence
Step 1- Students match relevant evidence to a given claim	Students are given all of the evidence and the claim. Minimal or no extra content knowledge is required	Students are given all of the evidence and the claim. However, to choose the match the evidence to the claim, they must either interpret the evidence or apply extra content knowledge	Students are given evidence and a claim, however, in order to match the evidence to the claim, they must interpret the data to apply additional content knowledge
Step 2- Students choose a relevant claim and construct a simple explanation based on given evidence (construction is scaffolded)	Students are given evidence, to choose the claim and construct the explanation, minimal or no additional knowledge or interpretation of evidence is required	Students are given evidence, but to choose a claim and construct the explanation, they must interpret the evidence and/or apply additional content knowledge	Students are given evidence, but to choose a claim and construct the explanation, they must interpret the evidence and apply additional content knowledge.
Step 3-Students construct a claim and explanation that justifies claim using relevant evidence (unscaffolded)	Students must construct a claim and explanation however, they need to bring minimal or no additional content knowledge to the task	Students must construct a claim and explanation that requires either interpretation or content knowledge	Students must construct a claim and explanation that requires the students to interpret evidence and apply additional content knowledge.

Reasoning and evidentiary argument



Types of reasoning
Inductive, Deductive
Quality of reasoning…



Implications for complexity and alignment evaluation of NGSS assessments

- It is useful to have more **general claims** (e.g., PLDs, assessment blueprints) for some purposes (e.g., public reporting), and **more specific claims** for other purposes (e.g., test development, validation)
- **More specific definitions of science task and reasoning characteristics**, (such as attempted by Shavelson, Baxter, Wilson, Mislevy, Dueschl, Songer, etc.) may be useful to those developing conceptions of **expertise** for the NGSS science assessments
  - Need to adapt some to the NGSS (e.g., CCCs), but very helpful in providing options for further defining SEP and SEPxDCI complexity characteristics
  - Need more on evaluating reasoning of the whole (purposeful claim-evidence)

My Central Messages

- We need to create more specific ***claims*** regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.
- The claims may be thought of as assertions about the development and application of expertise, along several dimensions.
  - The NGSS in particular require taking a stance about *transfer* and *generalization* because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.
- Researchers and developers have come up with ***examples*** of how science expertise can be further defined and assessed
- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

Classroom and large-scale

- Time scale for interpretation and action
- Instruction immediately in-process or at a summary point
- Curriculum spirals or builds knowledge/skills over time

Dimension	Description of type of growth
Independence	"The student was able to do similar tasks with less support"
Correctness	"The student mastered things s/he had incorrect previously"
Near transfer (similar examples)	"The student addressed new, similar examples of the same knowledge and skills"
Cognitive complexity of same knowledge/skills	"The student addressed more complex examples of the same knowledge and skills"
Fluidity & confidence	"The student became quicker and more confident in applying the same knowledge and skills"
Far transfer (application) of knowledge and skills	"The student can apply knowledge and skills to new, different situations, generating new combinations and solutions"
Expert knowledge representations & skills	"The student learned more sophisticated and refined knowledge, and applied skills more expertly"
Internal, creative roles, purposes, & values	"The student shifts roles to generate her/his own problems, purposes, according to own values consistent with discipline"

My Central Messages

- We need to create more specific claims regarding the NGSS to direct assessment development, interpretation, use, validation—and instruction.
- The claims may be thought of as assertions about the development and application of expertise, along several dimensions.
  - The NGSS in particular require taking a stance about transfer and generalization because the Framework is so expansive but the Performance Expectations (PEs) are so thinly representative.
- Researchers and developers have come up with examples of how science expertise can be further defined and assessed
- Claims about development and application of expertise can help ***clarify*** and ***direct*** coherent connections between ***classroom*** and ***large-scale*** assessments.

For more information:

Center for Assessment
www.nciea.org

Brian Gong
bgong@nciea.org