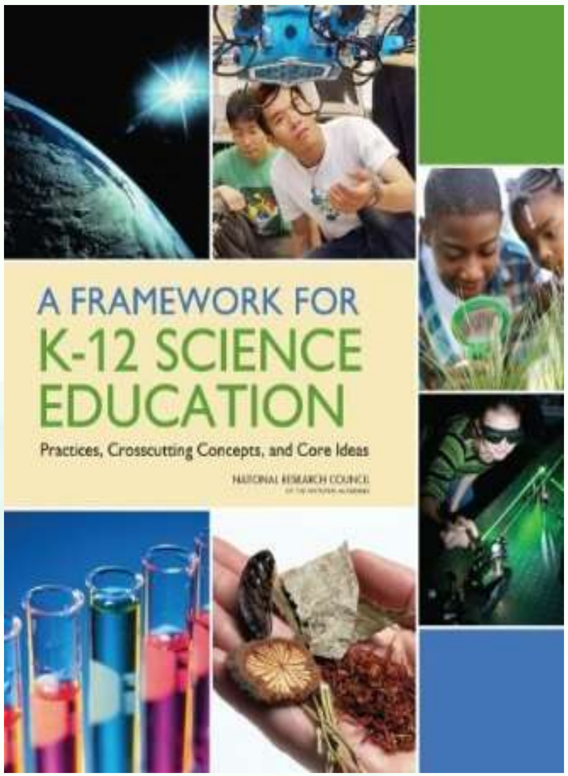
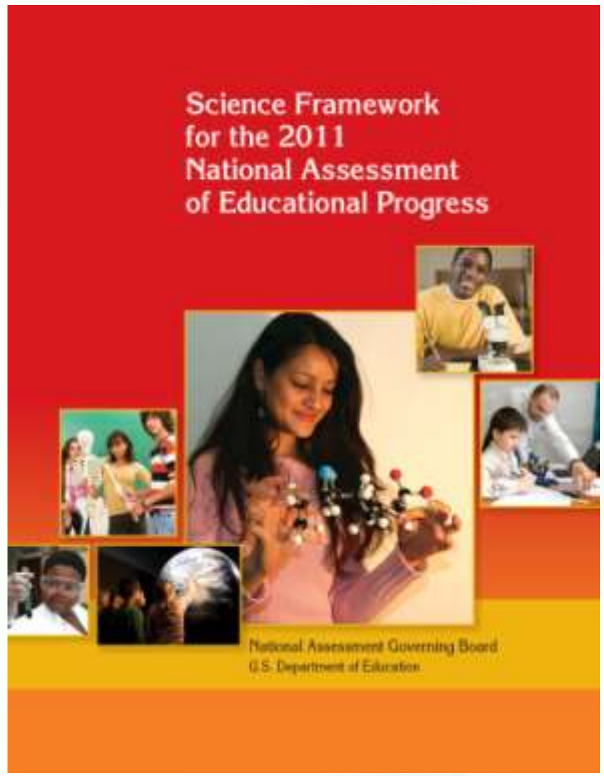
NAEP Assessments of Science Content Knowledge in Practice

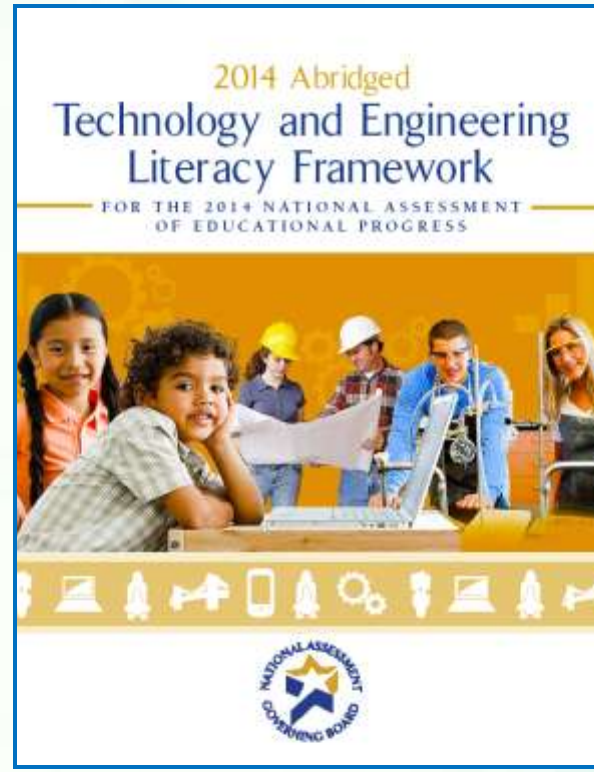
Hands-On and Interactive Computer Tasks

Dr. Alan J. Friedman
National Assessment Governing Board

Dr. Peggy G. Carr
National Center for Education Statistics


A FRAMEWORK FOR
K-12 SCIENCE
EDUCATION
Practices, Crosscutting Concepts, and Core Ideas
NATIONAL RESEARCH COUNCIL


Science Framework
for the 2011
National Assessment
of Educational Progress
National Assessment Governing Board
U.S. Department of Education

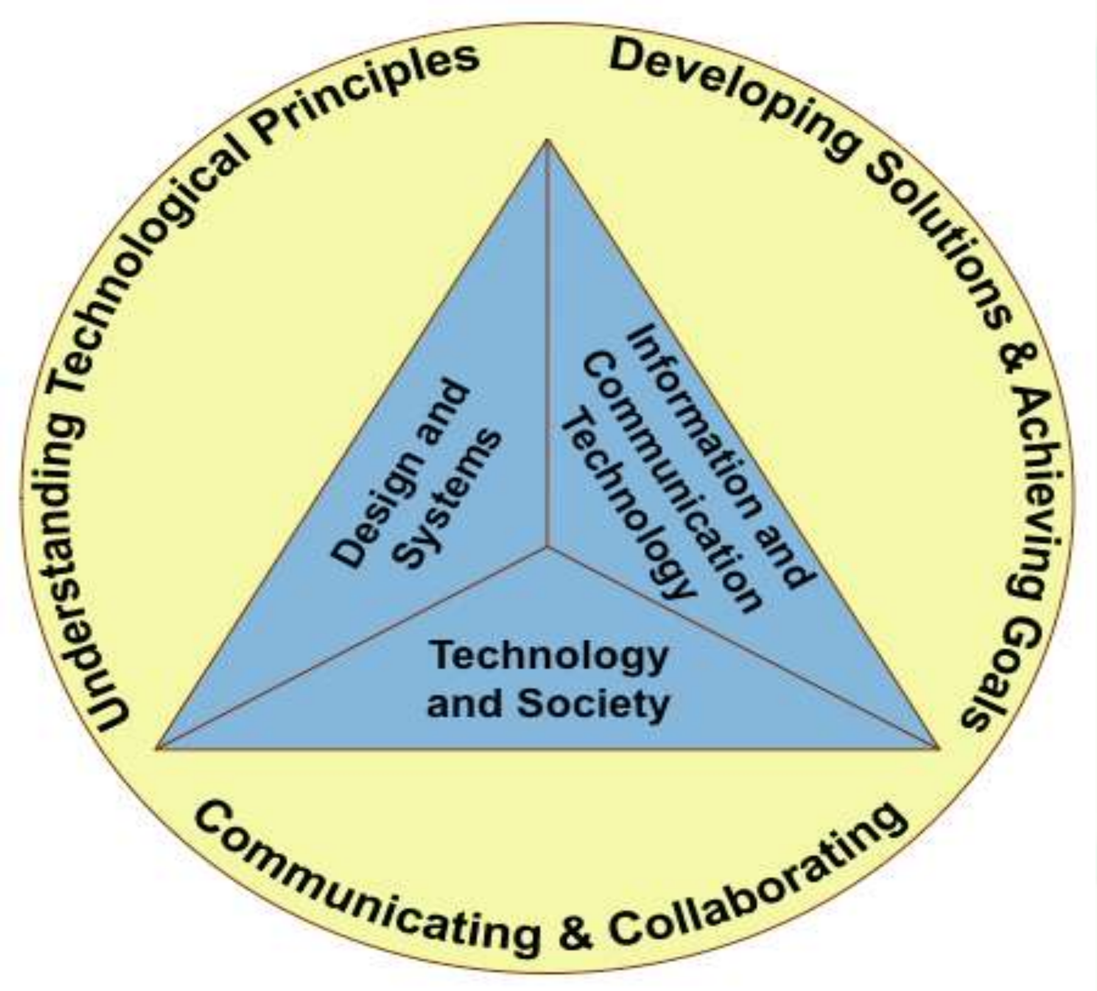

2014 Abridged
Technology and Engineering
Literacy Framework
FOR THE 2014 NATIONAL ASSESSMENT
OF EDUCATIONAL PROGRESS

2012 Framework for the K-12 Science Education Standards	NAEP 2009 Science Framework *NAEP 2014 Technology & Engineering Literacy (TEL) Framework*
"Scientific and Engineering Practices"	"Science Practices/*TEL Practices*"
Asking questions (for science) and defining problems (for engineering) Developing and using models Planning and carrying out investigations	Identifying Science Principles Using Science Principles *Understanding Technological Principles*
Analyzing and interpreting data Using mathematics and computational thinking Constructing explanations (for science) and designing solutions (for engineering)	Using Scientific Inquiry Using Technological Design *Developing Solutions and Achieving Goals*
Engaging in argument from evidence Obtaining, evaluating, and communicating information	*Communicating and Collaborating*

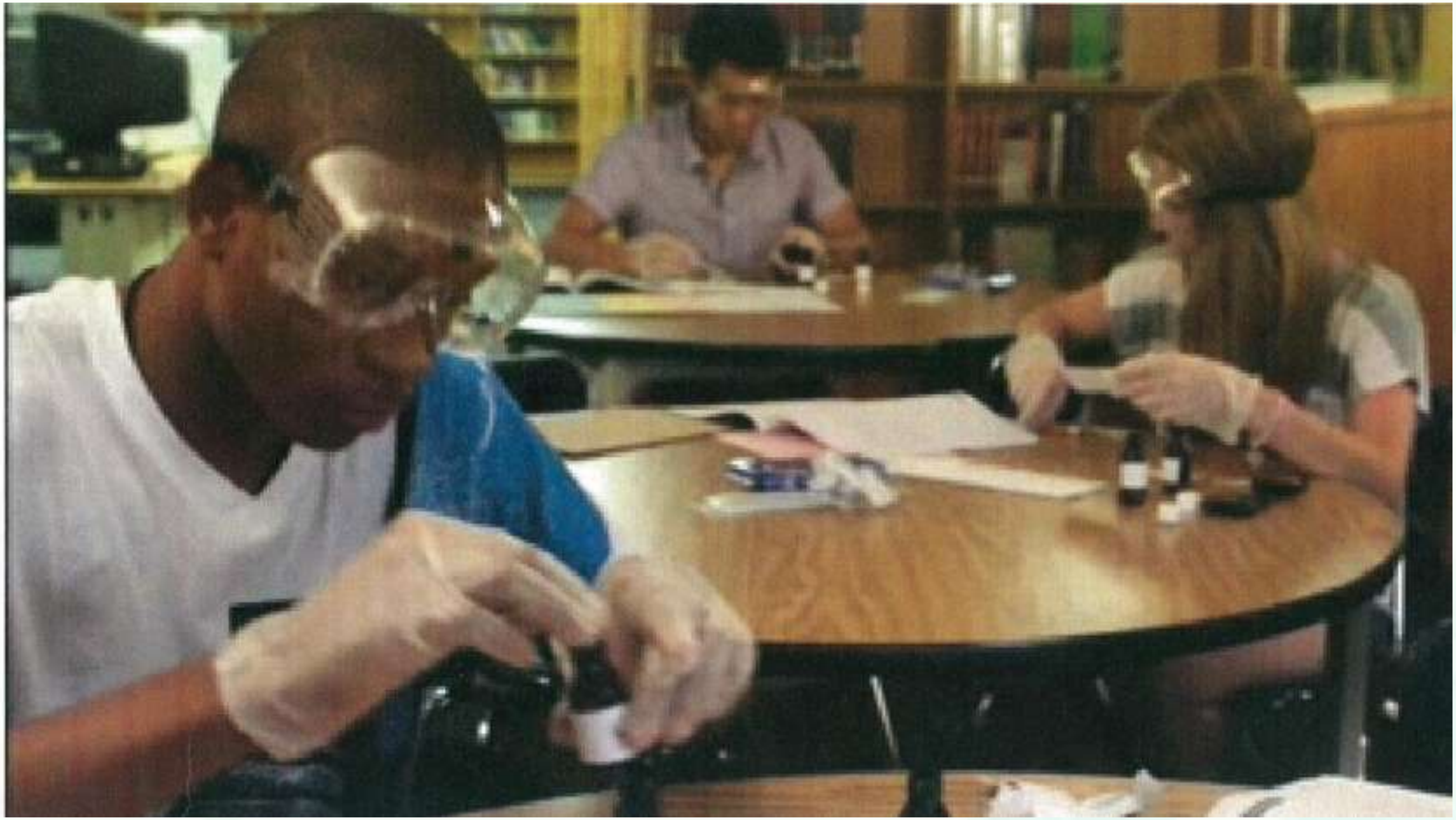
NAEP Components

- Multiple-choice items
- Short constructed response items
- Extended (long) constructed-response items
- Hands-on tasks
- Interactive computer tasks (ICTs)

NAEP Technology and Engineering Literacy (TEL) Assessment Framework

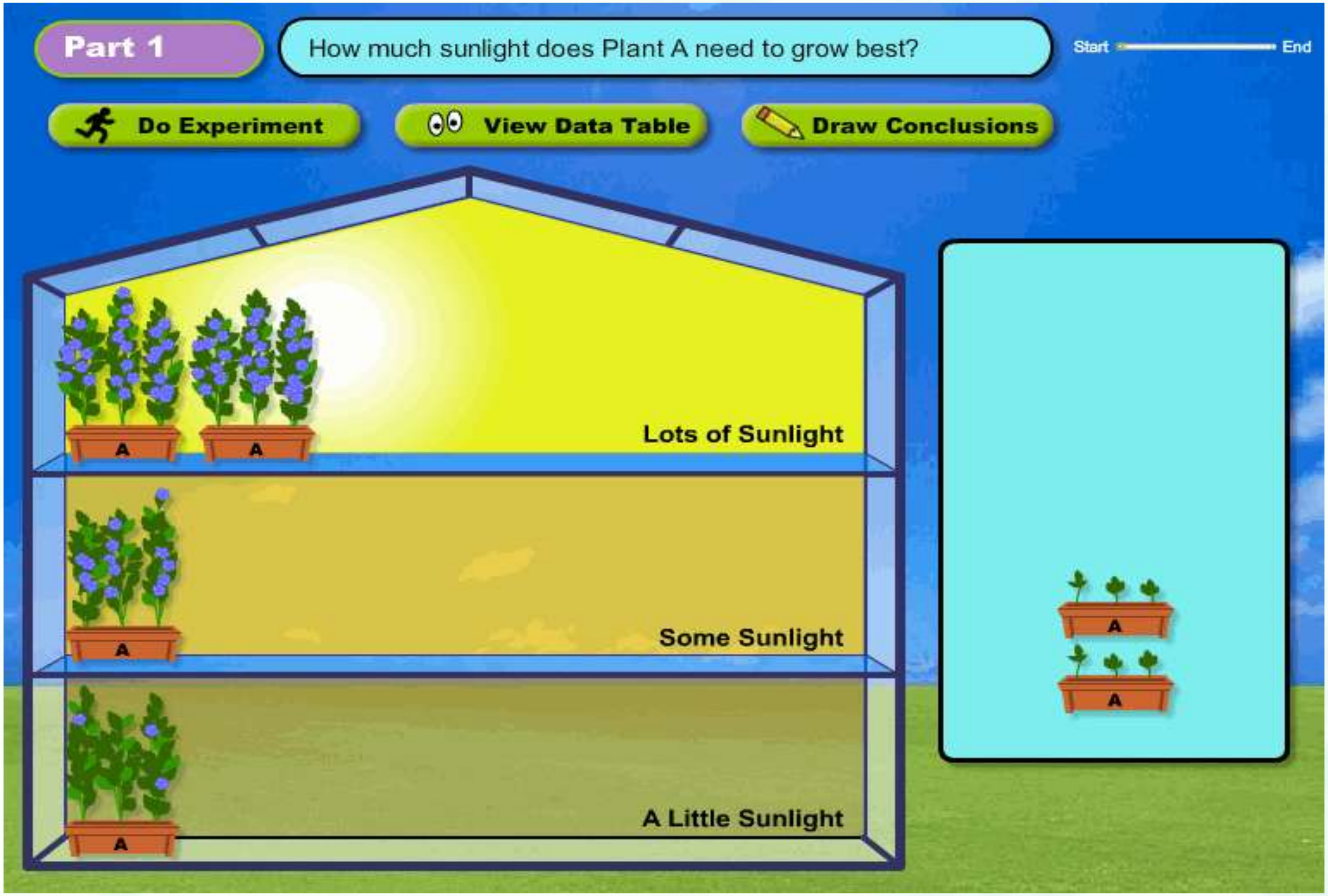


NAEP Hands-on Task (HOT)
Grade 12 – Maintaining Water Systems



NAEP Interactive Computer Task (ICT)

Grade 4 – Mystery Plants

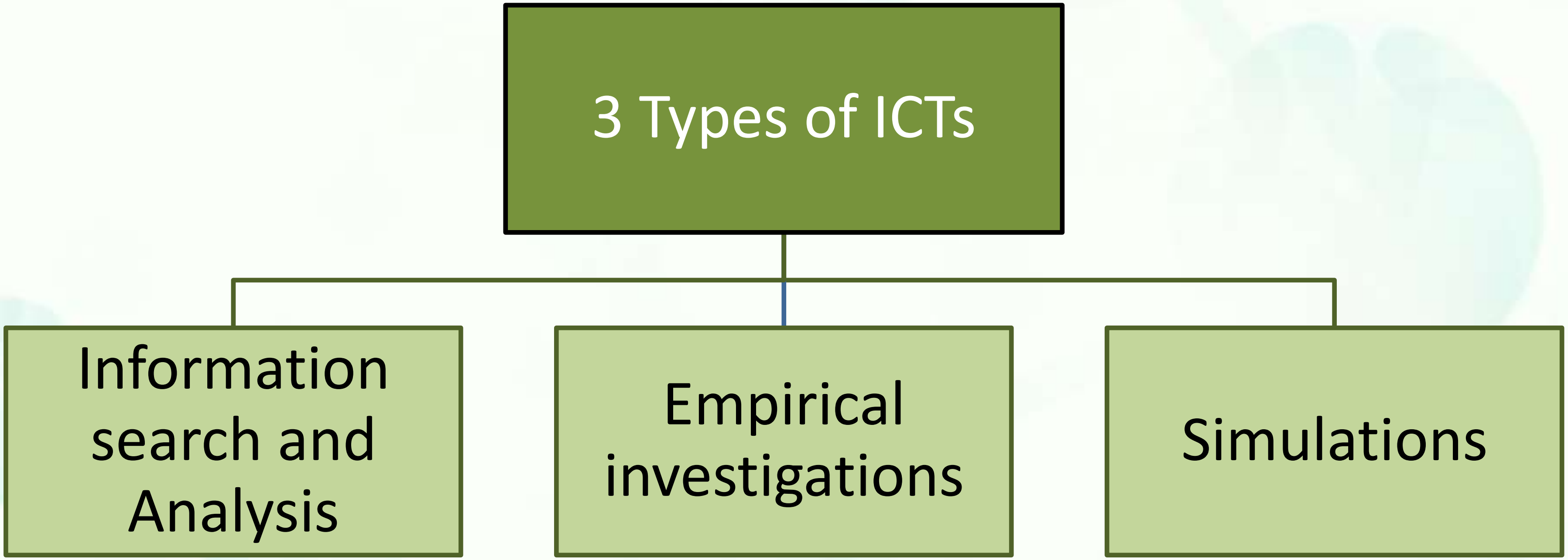


Interactive Computer Task Example

Advantages of ICTs

Studying phenomena that cannot be easily observed	• Observe plants grow • Observe microscopic organisms
Working safely in lab-like simulations	• Examining how heat relates to flow rates of liquids
Studying situations that require repetition	• Finding heat capacity of metals
Searching resources	• Use online documents to learn about previous studies on phytoplankton

Types of ICTs



Mystery Plants – Grade 4

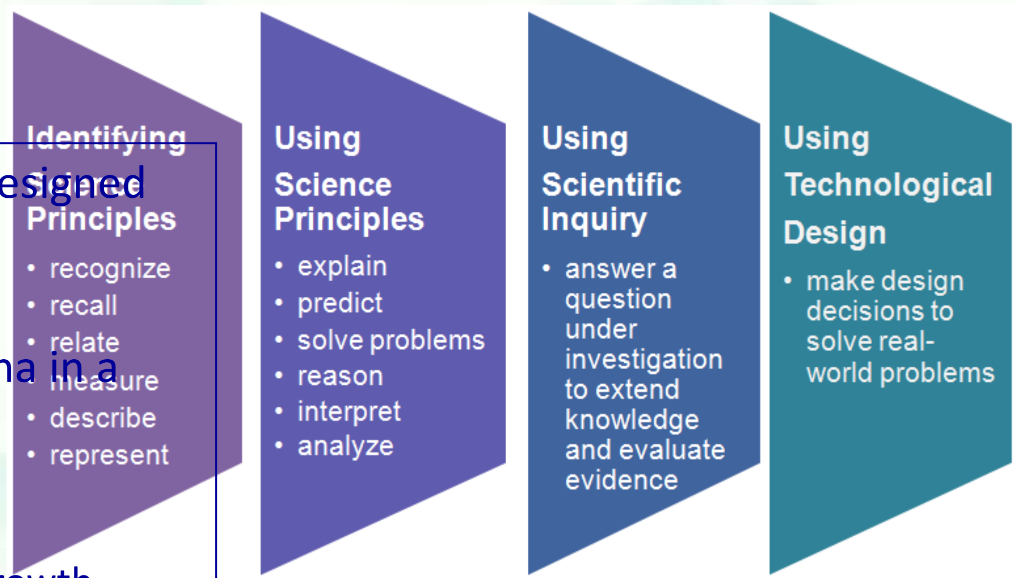
Science Content

Physical Science
- Matter
- Energy
- Motion

Life Science
- Structure and Functions of Living Systems
- Changes in Living Systems

Earth & Space Sciences
- Earth in Space and Time
- Earth Structures
- Earth Systems

Science Practices



- Empirical investigation of plant growth designed as an interactive computer simulation
- Students perform repetitious testing and observations of real biological phenomena in a simulated setting
- Key advantages of computer simulations
  - –Time can be sped up for observing plant growth
  - –Abstract ideas can be made tangible and accessible
  - –Simplified, real-world setting allows for focusing on construct relevant elements of the question, which reduces unnecessary cognitive actions
- Presented as a Predict-Observe-Explain problem set

Mystery Plants Grade 4

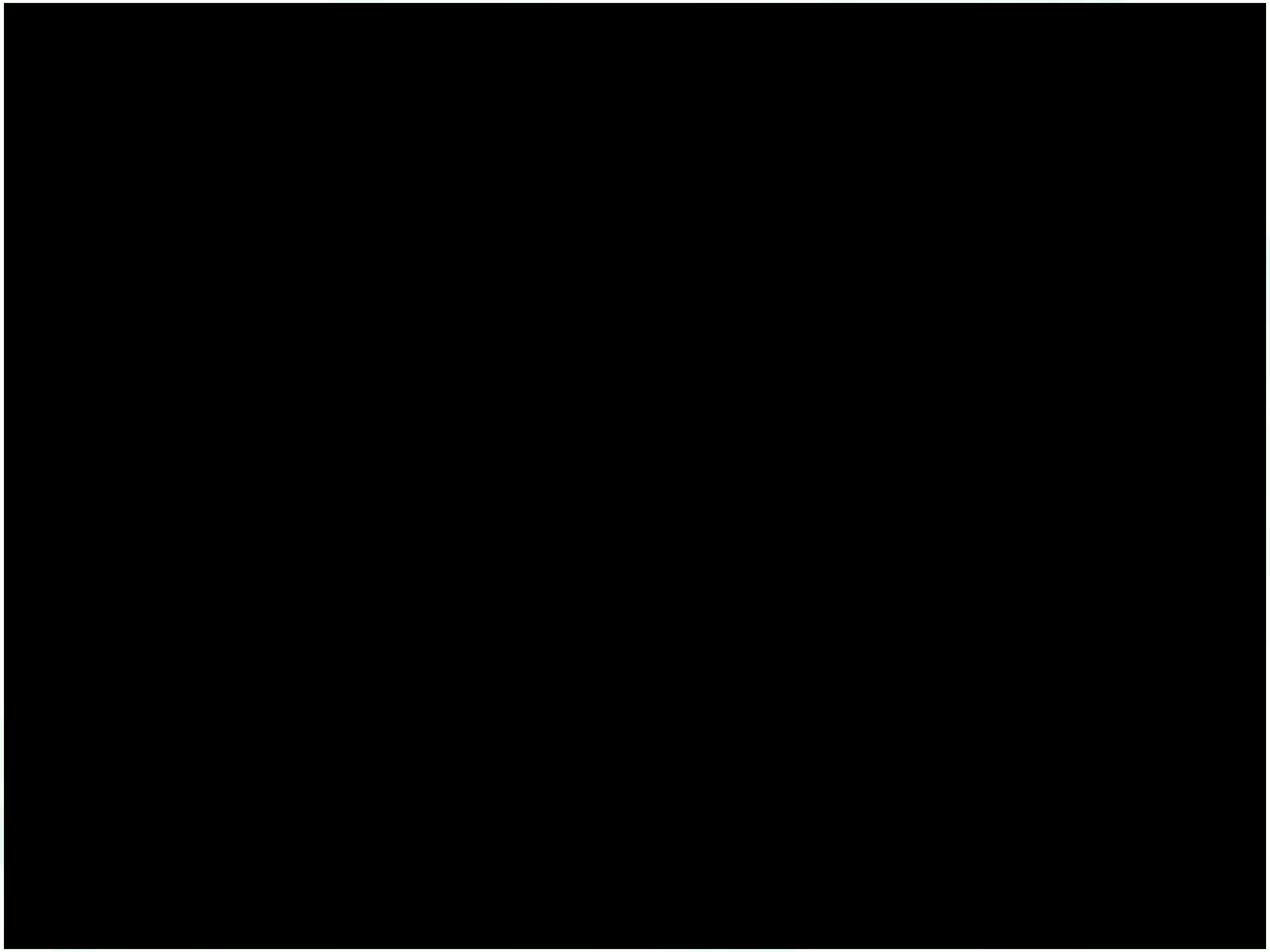
Experiment 1

What are the best sunlight conditions for growth of Plant A (sun-loving plant)?

Experiment 1: Sun-loving plant

Predict

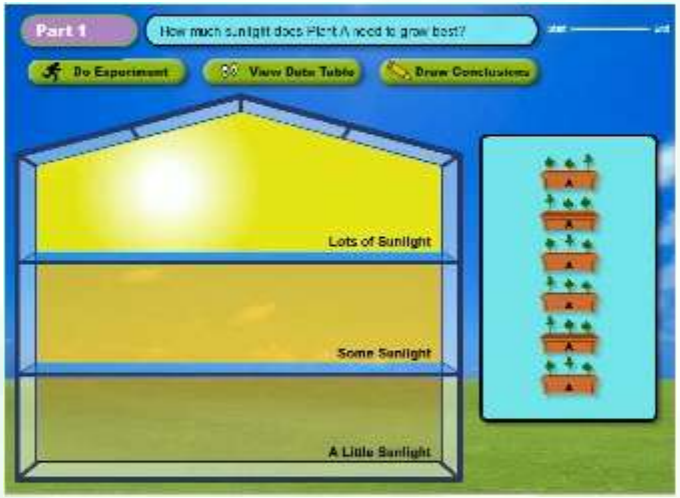
Use prior knowledge to predict the levels of sunlight plants need.



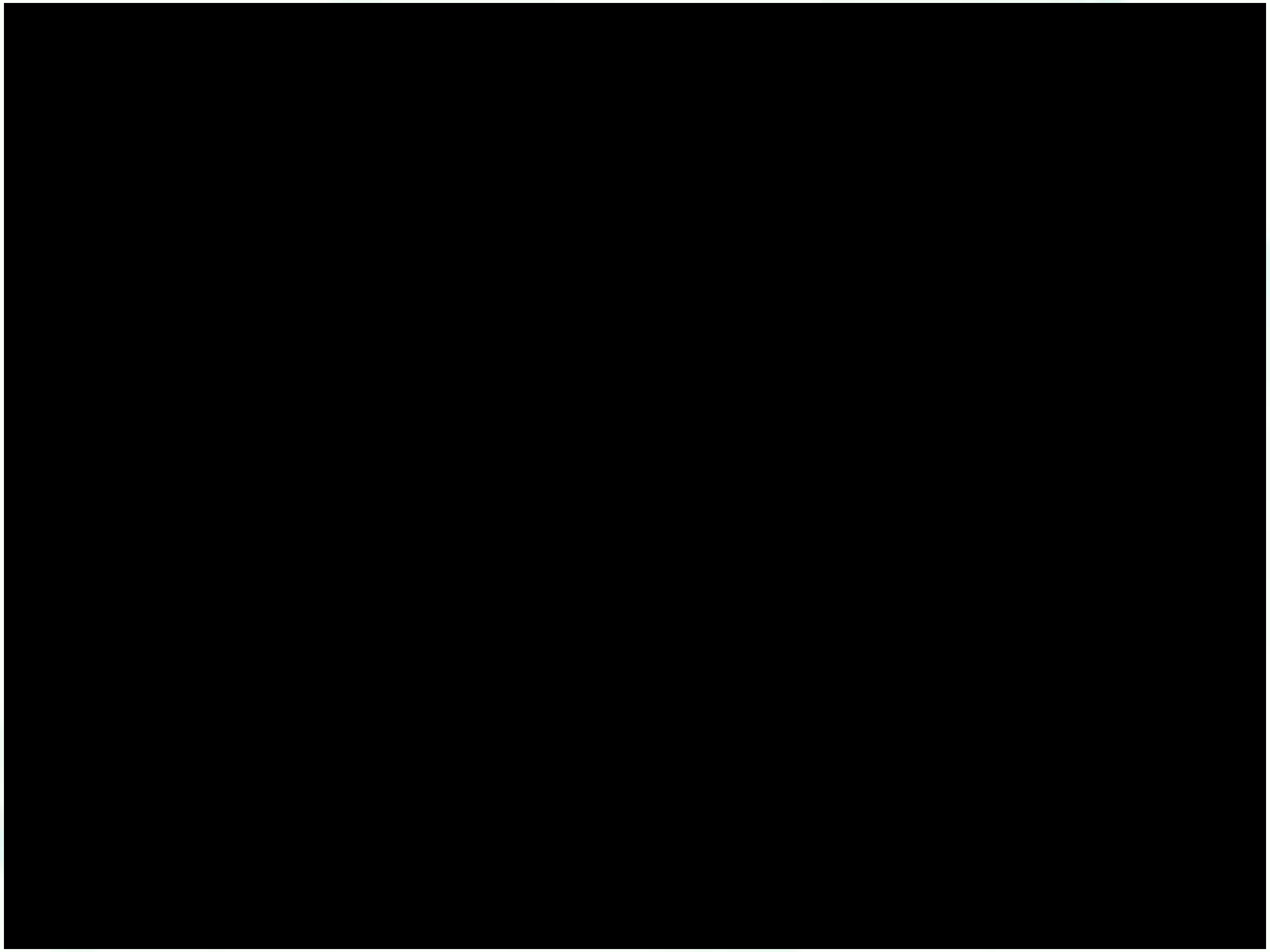
Prediction question assesses students' ability to make predictions by reasoning with scientific facts, concepts, and principles

Experiment 1: Sun-loving plant

Observe



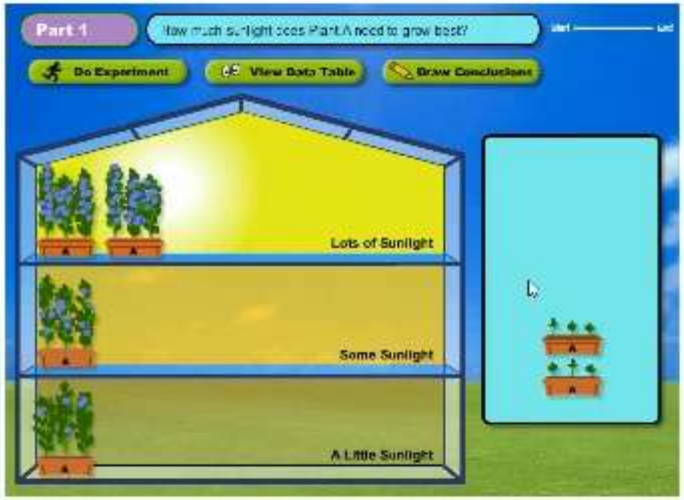
Investigate and make observations about how varying sunlight amounts affect the growth of Plant A.



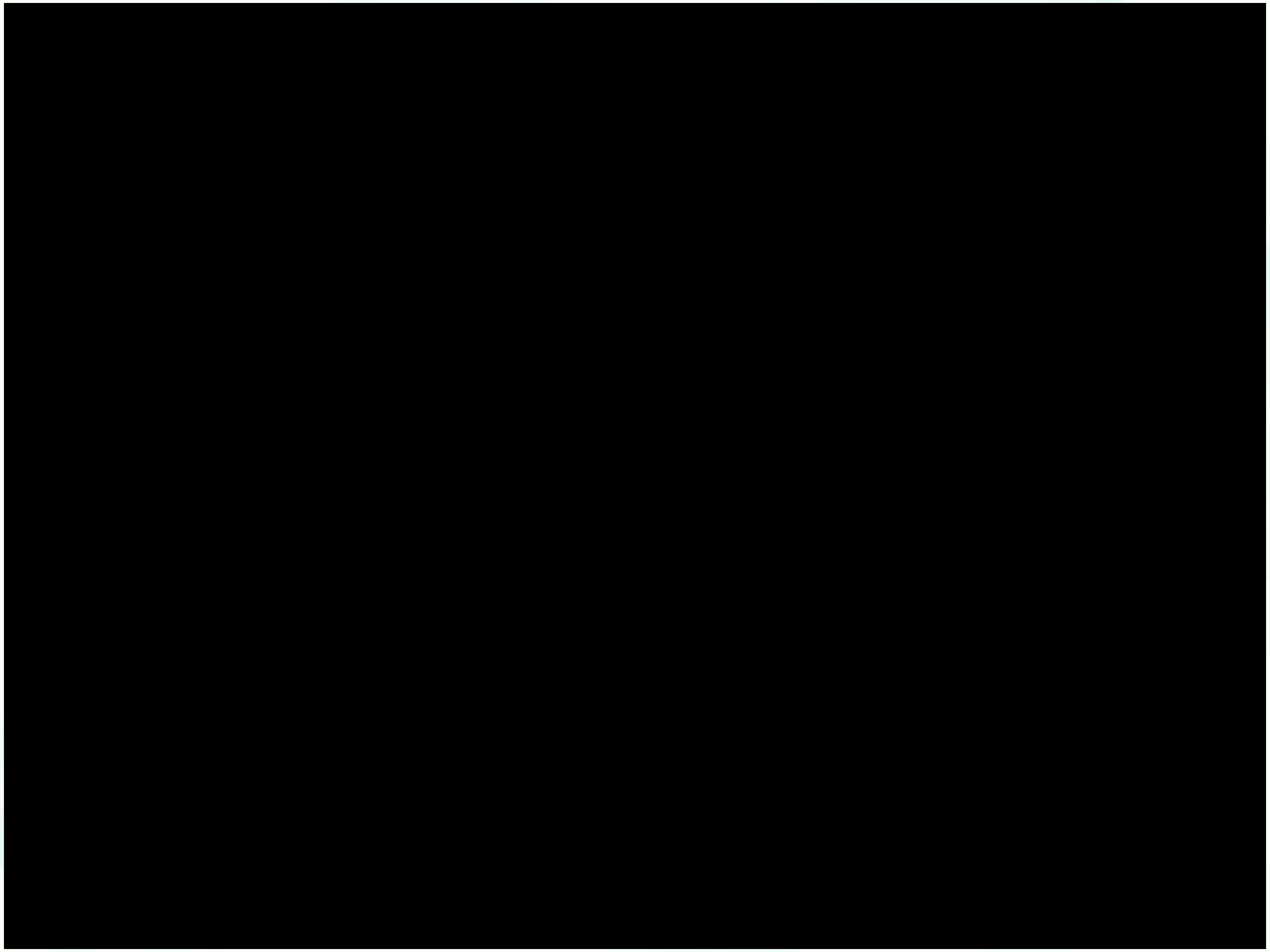
Observation task assesses students' ability to know how and why science proceeds as it does by using empirical testing

Experiment 1: Sun-loving plant

Explain



Select the correct conclusion and provide an explanation using data observed.



Explanation questions assess students' ability to explain how the natural world works

How the Assessment Is Administered

Administering the Assessment

- Administered by NAEP field staff
- Field staff bring in all necessary equipment (test booklets, hands-on tasks, laptops, ear buds, etc.)
- Staff are responsible for all pre-assessment and assessment day activities
- Students are assessed in classrooms or as a group in a school cafeteria or other large room

Assessment Design

Interactive Computer Tasks	Sample
• Two 20-minute tasks • One 40-minute task	• Grades 4, 8, and 12 • 2,000 students per grade

Scoring Tasks and Reporting Scores

Scoring Tasks

- Scored similarly to all other NAEP assessments
  - Scorers evaluated constructed-response items according to scoring rubrics
  - Multiple-choice answers were machine-scored
- Collected student actions in extended ICTs
  - Examined how well students used computer-based tools to conduct scientific investigations

Reporting Scores

- Percent correct
  - Performance was summarized at the item level by using the average percentage correct
- Student percent correct score
  - Performance was summarized across test items using a student percent correct score
- Process analyses
  - Grouped students into various categories according to their response patterns to a pre-specified item sequence

Examination of Percent Correct Patterns

Key Discoveries

1

Students were **successful** on parts of investigations that involved limited sets of data and making straightforward observations of that data.

2

Students were **challenged** by parts of investigations that contained more variables to manipulate or involved strategic decision making to collect appropriate data.

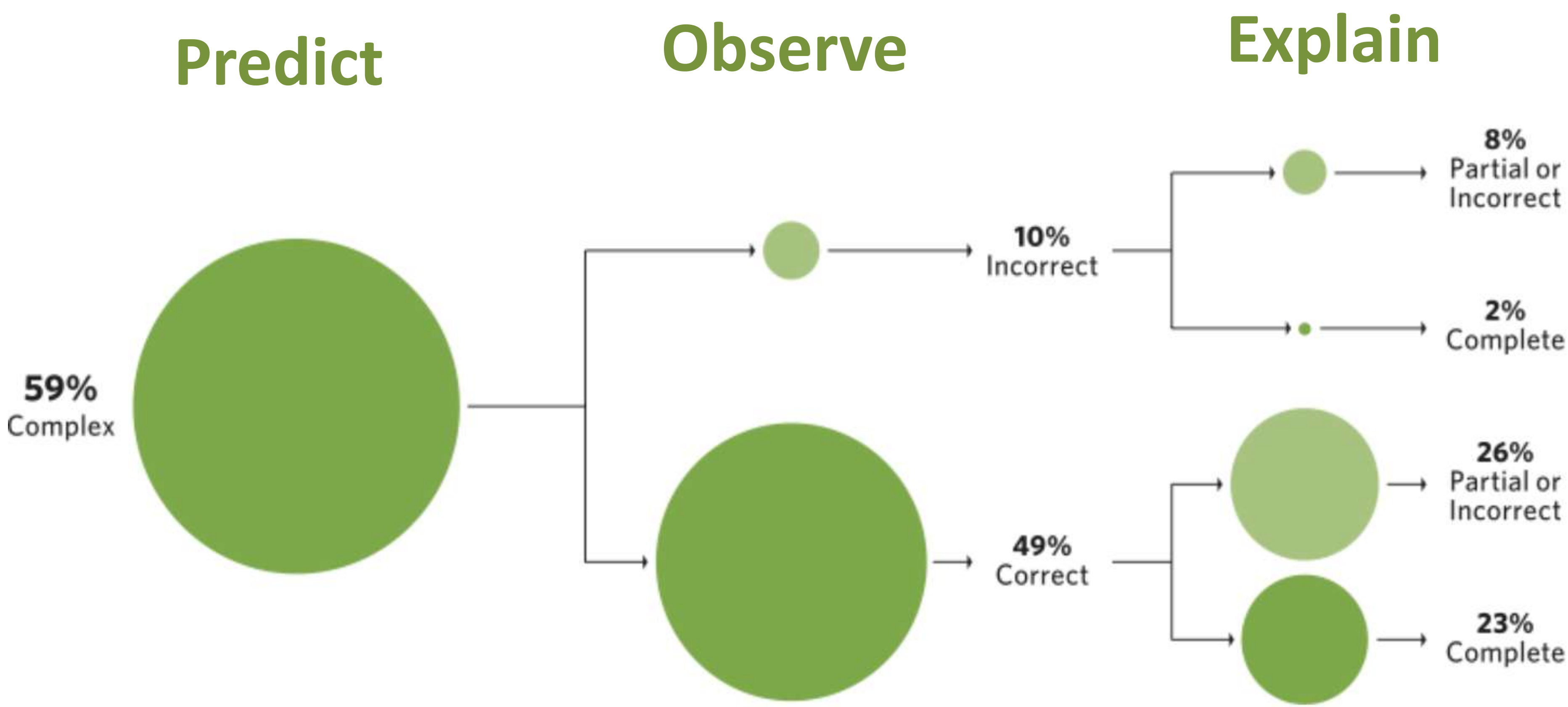
3

Students could select **correct** conclusions from an investigation, but scored **lower** when asked to explain their results.

Reporting on Process

- **Practices, as defined by the science framework, have not been reported on separately in the NAEP assessments**
- **Analysis of science HOTs and ICTs focuses on processes and captures solution paths for tasks**
  - The 2009 examples show specific areas of strength and weakness in student success across a task

Only 23% of all fourth-graders displayed complex prior knowledge, did the experiment correctly, and were able to give complete explanations. (Follow the leftmost series of green disks)



Practical Issues, Challenges, and Innovations

Development Timeline

- Assessment development is a multi-year process

Fall 2006: Begin pilot development

Winter 2008: Perform task try-outs and stakeholder reviews

Spring 2008: Administer pilot and score response

Summer 2008: Perform analyses and select items for operational assessment

Winter-Fall 2009: Administer assessment, score responses, and analyze data

Spring 2010 - 2011: Analyze content for reporting and produce reports

Task Development

- **Challenge**: Costly to design and build
- **Solutions**: Early steps ensure accurate measure
  - Staged approach to development:
    - Student feedback in tryouts
    - Performance on pilot
  - Balance between engagement and measurement
  - Collaboration with team of developers, designers, cognitive scientists, and psychometricians

Task Development

- **Challenge**: Create valid, reliable, and developmentally appropriate tasks
- **Solutions**:
  - Multiple tryouts with students
    - Less than 10 interviewed about storyboard
    - Small-scale tryout with 50 students in prototype format
    - Large-scale tryout with 300 students (national sample) close to final form, measured performance and ability to use CBA tools

Task Development

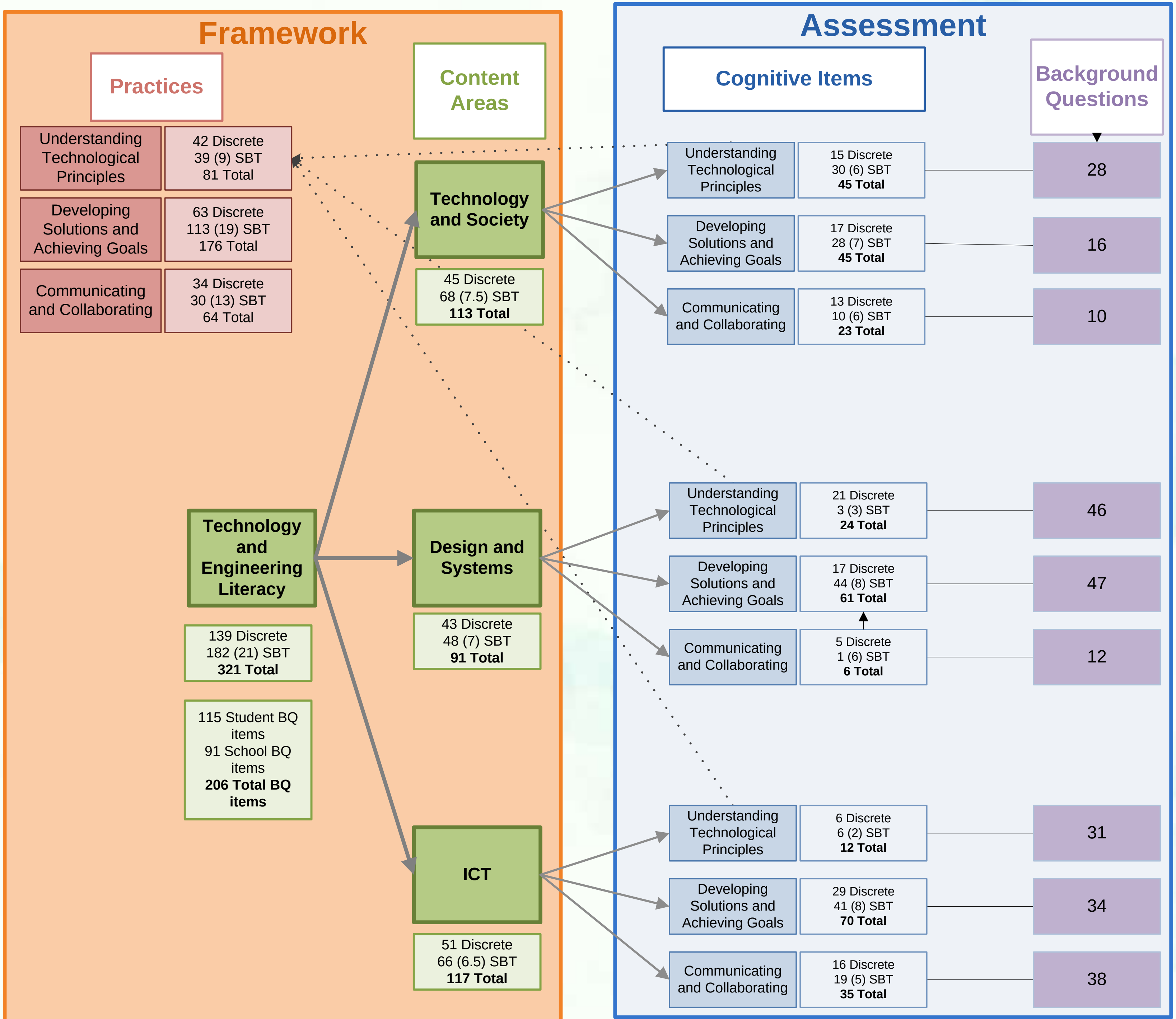
- **Challenge**: Provide environment to support student success
- **Solutions**:
  - Create multi-faceted tutorial about interface and frame the science problem
  - Make “help” buttons clear for additional support
  - Explicit introductory and section directions
  - Welcome & thank you maintains the scenario

Analysis & Scaling Challenges

- An IRT scale score was not reported for the Science ICTs because there were too few tasks
  - Three times as many tasks are being developed to support a scale in 2015
- Scenario-based complex tasks often invoke particular dependencies between items
  - Other factors besides proficiency may cause correlations across items

Innovations

- Use of Evidence-Centered Design
  - Technology and engineering literacy
  - Next-generation science ICTs
  - Other assessments



Framework
Practices
Content Areas
Understanding Technological Principles
42 Discrete
39 (9) SBT
81 Total
Developing Solutions and Achieving Goals
63 Discrete
113 (19) SBT
176 Total
Communicating and Collaborating
34 Discrete
30 (13) SBT
64 Total
Technology and Society
45 Discrete
68 (7.5) SBT
113 Total
Technology and Engineering Literacy
139 Discrete
182 (21) SBT
321 Total
115 Student BQ items
91 School BQ items
206 Total BQ items
Design and Systems
43 Discrete
48 (7) SBT
91 Total
ICT
51 Discrete
66 (6.5) SBT
117 Total
Assessment
Cognitive Items
Background Questions
Understanding Technological Principles
15 Discrete
30 (6) SBT
45 Total
28
Developing Solutions and Achieving Goals
17 Discrete
28 (7) SBT
45 Total
16
Communicating and Collaborating
13 Discrete
10 (6) SBT
23 Total
10
Understanding Technological Principles
21 Discrete
3 (3) SBT
24 Total
46
Developing Solutions and Achieving Goals
17 Discrete
44 (8) SBT
61 Total
47
Communicating and Collaborating
5 Discrete
1 (6) SBT
6 Total
12
Understanding Technological Principles
6 Discrete
6 (2) SBT
12 Total
31
Developing Solutions and Achieving Goals
29 Discrete
41 (8) SBT
70 Total
34
Communicating and Collaborating
16 Discrete
19 (5) SBT
35 Total
38