Assessing Computational Thinking

Eric Snow, SRI International
Irvin Katz, Educational Testing Service
Allison Elliott Tew, University of British Columbia
Jill Feldman, Westat

NSF-CE21 Community Meeting
Washington, DC
February 2-3, 2012

Theme 1

Is "Grandma Baking a Cake" Computational Thinking (CT)?

Probably not, but there is emerging consensus on certain characteristics of CT, *even if their importance, specification and scope vary across use contexts*.

Progress is being made across multiple contexts in identifying and measuring these characteristics.

Sessions

Session 1: Review of CT Measurement Domain, Key Assessment and Evaluation Efforts to Date

- Activity: Important Proficiencies in the CT Domain

  Guest Panelist: Allison Elliot Tew, UBC

Theme 2

Certain "Myths" Persist About Assessment Design and Validity.

They can influence how assessments are talked about in proposals, as well as how they are designed and developed in practice.

They become particularly appealing when working in novel measurement domains, such as CT, where the characteristics of the domain are still emerging, and only a few assessments exist.

Sessions

Session 1: Review of CT Measurement Domain, Key Assessment and Evaluation Efforts to Date

- Activity: Important Proficiencies in the CT Domain

  Guest Panelist: Allison Elliot Tew, UBC

Session 2: Validity, Mythology and Assessment Design

- Activity: Identifying and Aligning Evidence with Focal CT Proficiencies

  Guest Panelist: Irvin Katz, Educational Testing Service

Theme 3

Evaluation is a Critical Component of Competitive NSF Proposals.

Theories of Action and *Logic Models* can help investigators get organized, and specify and align important study components (e.g., inputs, activities, outcomes).

Assessment plays a critical role in most evaluations and needs to be considered carefully before being used to measure short- and long-term program outcomes.

Sessions

Session 1: Review of CT Measurement Domain, Key Assessment and Evaluation Efforts to Date

- Activity: Important Proficiencies in the CT Domain

  Guest Panelist: Allison Elliot Tew, UBC

Session 2: Validity, Mythology and Assessment Design

- Activity: Identifying and Aligning Evidence with Focal CT Proficiencies

  Guest Panelist: Irvin Katz, Educational Testing Service

Session 3: Program Evaluation

- Activity: Tasks/Situations for Eliciting Evidence of Focal CT Proficiencies

  Guest Panelist: Jill Feldman, Westat

A Note About Session Activities

We want to give you an opportunity to tell us about the importance of CT proficiencies, types of evidence for those proficiencies, and situations/tasks that might elicit that evidence.

Please complete the activity worksheets as clearly as possible and hand them in when you are finished with each activity.

Your responses will remain anonymous and any results will be reported in aggregate; we hope to share results following the workshop.

Review of CT Measurement Domain, Key Assessment & Evaluation Efforts to Date

Allison Elliott Tew, University of British Columbia

What is Computational Thinking?

- Term first used by Seymour Papert (1996)
  - *"In both cases the computer used as a tool effectively leads to a solution, but in neither does the computational representation make the mathematics more perspicuous. … The goal is to use computational thinking to forge ideas that are at least as 'explicative' as the Euclid-like constructions (and hopefully more so) but more accessible and more powerful."*

What is Computational Thinking?

- Popularized by Jeanette Wing in ACM Viewpoints Article

"I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it…"

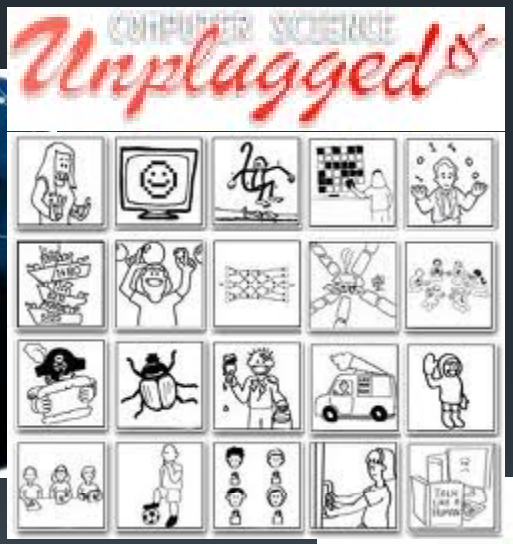
- Justice Potter Stewart (1964)

- Ways to Think Like a Computer Scientist

Examples of CT

App Inventor

ECS Exploring Computer Science



Operationalizing CT (AP)

- CS Principles Computational Thinking Practices (College Board, 2011)
  1. Connecting computing
  2. Developing computational artifacts
  3. Abstracting
  4. Analyzing problems and artifacts
  5. Communicating
  6. Working effectively in teams
- Sample Population
  - Introductory university course

Operationalizing CT (AP)

- CS Principles Computational Thinking Practices (College Board, 2011)
  1. Connecting Computing
     a. Identification of impacts of computing.
     b. Description of connections between people and computing.
     c. Explanation of connections between computing concepts.
  2. Developing computational artifacts
     a. Creation of an artifact with a practical, personal, or societal intent.
     b. Selection of appropriate techniques to develop a computational artifact.
     c. Use of appropriate algorithmic and information-management principles.
  3. Abstracting
     a. Explanation of how data, information, or knowledge are represented for computational use.
     b. Explanation of how abstractions are used in computation or modeling.
     c. Identification of abstractions.
     d. Description of modeling in a computational context.
  4. Analyzing problems and artifacts
     a. Evaluation of a proposed solution to a problem.
     b. Location and correction of errors.
     c. Explanation of how an artifact functions.
     d. Justification of appropriateness and correctness.
  5. Communicating
     a. Explanation of the meaning of a result in context.
     b. Description using accurate and precise language, notation, or visualizations.
     c. Summary of purpose.
  6. Working effectively in teams
     a. Application of effective teamwork practices.
     b. Collaboration of participants.
     c. Production of artifacts that depend on active contributions from multiple participants.

Operationalizing CT (K-12)

- International Society for Technology in Education (ISTE) and CS Teacher's Association (CSTA) (2011)
- A problem-solving process that includes (but is not limited to) the following characteristics:
  - Formulating problems in a way that enables us to use a computer and other tools to help solve them
  - Logically organizing and analyzing data
  - Representing data through abstractions such as models and simulations
  - Automating solutions through algorithmic thinking (a series of ordered steps)
  - Identifying, analyzing, and implementing possible solutions with the goal of achieving the most efficient and effective combination of steps and resources
  - Generalizing and transferring this problem-solving process to a wide variety of problems
- Dispositions or attitudes that are essential dimensions of CT
  - Confidence in dealing with complexity
  - Persistence in working with difficult problems
  - Tolerance for ambiguity
  - The ability to deal with open-ended problems
  - The ability to communicate and work with others to achieve a common goal or solution

Summary

- CT is a broad domain
- Everyone has an idea or theory about what is/isn't CT
- Assessments (& grants) don't need to wait for agreed upon standards
- Clearly define the knowledge, skills and attitudes (KSA) of CT in your learning activities and environments

Validated Assessments in CS

Validated Exams	CS Education Research Efforts
High School Level Exams	SUNY CS1-2 Exam
Advanced Placement (AP)	(Decker, 2007)
General Certificate of Education (A-Levels)	
CS1	CS Concept Inventory
FCS1 Assessment (Elliott Tew, 2010)	(Herman, Loui, & Zilles, 2010)
College Program Completion	Principled Assessment for Computational Thinking (PACT)
Major Field Test	CE21 Planning Grant (2011-2012)
GRE Subject Test	

- Challenges
  - Programming Language
  - Conceptual vs. syntactic knowledge
  - Agreement on content

FCS1 Assessment Instrument

- Measures a set of foundational CS1 concepts that are common across a wide variety of current pedagogical approaches and paradigms
  - E.g., Variables, operators, loops, arrays, recursion
- Psuedo-code used for programming language independence
  - Think aloud Interview (n = 13) and empirical studies (n = 952) confirmed that students are able to express their understanding of FCS1 concepts in a language independent exam
- Multiple choice question item format
  - Definition
  - Tracing
  - Code Completion
- Validated measure of introductory computing concepts for procedurally-based introductory computing courses taught in Java, Matlab, or Python at the university level.

Interested in the FCS1?

- Sign-up:

  http://tinyurl.com/4v8ktow

- Contact Information:

  aetew@cs.ubc.ca

- Please note:
  - ***This is an ongoing research project.***
  - ***We are developing an online version to facilitate dissemination and data collection.***
  - ***Feel free to send comments and feedback.***

Supported in part by the National Science Foundation

CISE #0306050

CCLI-ASA #0512213

CPATH CB #0829601

BPC-A #0634629

Principled Assessment for Computational Thinking (PACT)

CE21 Planning Grant (9/2011 – 4/2012)

Principal Investigators: Eric Snow, Marie Bienkowski, SRI International

Project Goal

Leverage interdisciplinary expertise to begin creating an assessment framework and design templates for computational thinking (CT) aligned with the *Exploring Computer Science* curriculum.

Principled Assessment for Computational Thinking (PACT)

Research Questions

1. In what ways are the organizing concepts and principles for the computational thinking domain represented in the *Exploring Computer Science* (ECS) curriculum?

1. What are the broad focal knowledge, skills and attitudes (KSAs) underlying each of the CT organizing concepts and principles represented in the ECS curriculum?

1. How should the organizing concepts and principles and underlying broad focal KSAs be organized into a structured framework and design templates that can support the assessment of computational thinking?

2. How can the assessment framework and design templates be used to support the assessment of computational thinking outcomes aligned with the ECS curriculum?

Principled Assessment for Computational Thinking (PACT)

Project Plan

December 201

Core expert group meets to begin creating an assessment framework and design templates for computational thinking (CT) aligned with the *Exploring Computer Science* curriculum

January 2012

Debrief core expert group meeting and circulate draft framework and design templates for internal review. Develop project dissemination web site and draft white paper on CT assessment. Begin planning for next stage of CT assessment design and development work

February – April 2012

Release project dissemination website, invite CE21 community feedback on assessment framework, design templates and white paper. Plan for next stage of CT assessment design and development work

Principled Assessment for Computational Thinking (PACT)

Web Site URL:

http://pact.sri.com

Please note:

- ***The site is still under construction.***
- ***PACT materials will be posted over coming 2 months.***
- ***Feel free to send comments and feedback.***

Questions & Comments

Now it's your turn. Please feel free to ask questions and comment on your work in this area.

Also consider:

- What are some of the important CT proficiencies across different contexts?
- To what extent are these proficiencies represented in the existing frameworks and standards documents?
- What about emerging CT proficiencies and attempts to assess them?

Activity 1
Important Proficiencies in the CT Domain

We want to hear from people in the field:

> *What are some of the important knowledge, skills and attitudes underlying CT proficiencies?*

Please review the Activity 1 worksheet.

Work in small groups, or individually, to rate the importance of a subset of CSTA standards.

We encourage you to consider and discuss the contexts in which you work as you rate the importance of the standards.

Validity, Mythology, and Assessment Design

Irvin Katz, Educational Testing Service

Validity

"…degree to which evidence and theory support the interpretations of test scores [in the context of proposed test uses]."

Standards for Educational and Psychological Testing, pg 9 – 1999: AERA, APA, NCME

Myth #1

"If I can find a validated assessment,
I'll be ready to evaluate my project."

Reality #1

Assessments are **not** "plug and play"

Need to check compatibility…

Knowledge, skills,
and attitudes (KSAs)

My Instruction

Evidence of KSAs

Their Assessment

Tasks that elicit
evidence

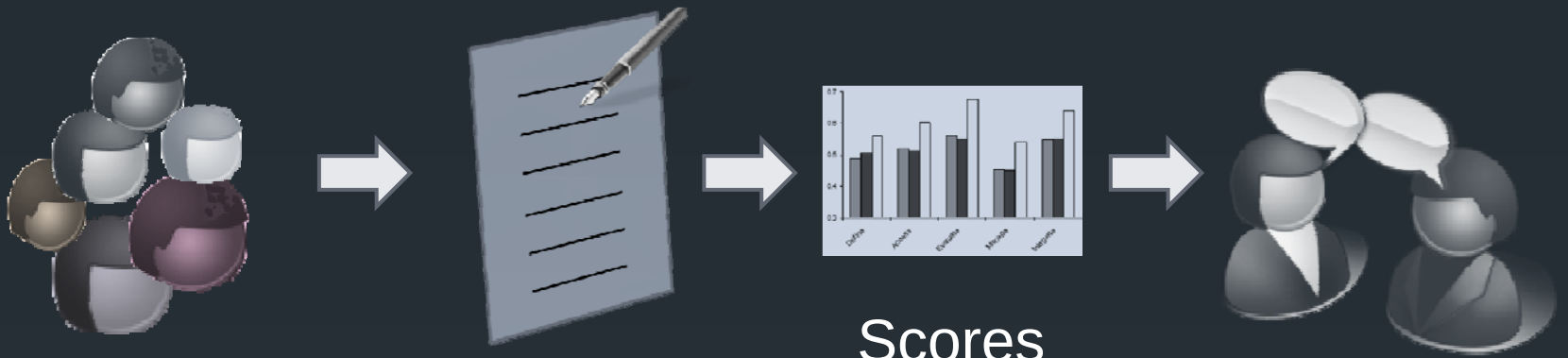
Myth #2

"Once an assessment is validated,
I can use it for my project."

Reality #2

Validity is **use-specific**

“Validated”



Students

Assessment

Scores

Decisions

KSAs

Evidence

Tasks

Design

last

Myth #3

"Assessment is all about having great tasks."

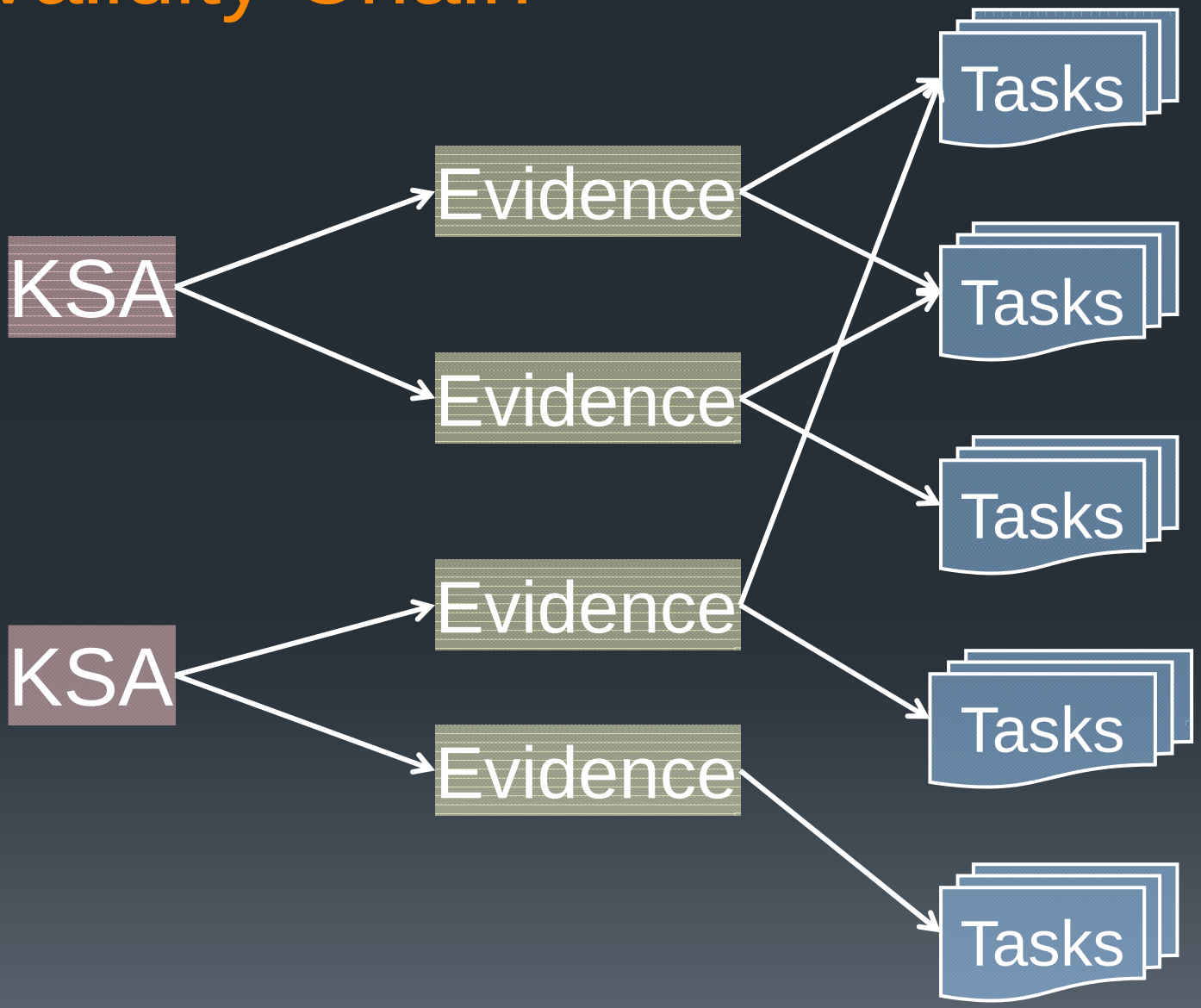
Reality #3

Assessment design **focuses on evidence**

Assessment Design

- What KSAs do I want to assess?
- What would be evidence of those KSAs?
- What tasks would elicit the correct evidence?

Validity Chain



Pop Quiz

Reality #1

Assessments are **not** "plug and play"

Reality #2

Validity is **use-specific**

Reality #3

Assessment design **focuses on evidence**

Realities of Assessments

- Assessments are not plug and play
- Validity is use-specific
- Assessment design focuses on evidence

Realities of Assessments

- Assessments are not plug and play
- Validity is use-specific
- Assessment design focuses on evidence

Realities of Assessments

- Assessments are not Plug and play
- Validity is Use-specific
- Assessment design focuses on Evidence

Questions?

Realities of Assessments

- Assessments are not plug and play
- Validity is use-specific
- Assessment design focuses on evidence

For more information:

http://www.ets.org/Media/Research/pdf/RD_Connections9.pdf

Questions & Comments

Now it's your turn. Please feel free to ask questions and comment on your work in this area.

Also consider:

- If assessments are not plug and play, and validity is use specific, then how, when working in a new domain like CT, should we address the assessment needs in our projects?

Activity 2
Identifying and Aligning Evidence with Focal CT Proficiencies

We want to hear from people in the field:

> *What counts as **evidence** of important CT proficiencies?*

Please review the Activity 2 worksheet.

Work in small groups, or individually, to identify possible types of evidence for a subset of CSTA standards.

We encourage you to consider and discuss the contexts in which you work as you identify possible types of evidence.

Program Evaluation

Jill Feldman, Westat

Session Objectives

- Provide practice-based tips for designing evaluation plans that increase proposal coherence
- Suggest ways of using graphic organizers that highlight key elements of NSF (and other funders') solicitations
- Critically reflect on the role of assessment in evaluation

How Does NSF Define "Educational Research*?"

Educational research…

- solves a problem
- requires familiarity with the field; competence in methodology; technical skill in collecting and analyzing data
- is based upon observable experience or empirical evidence
- demands accurate description and carefully documentation
- employs carefully designed procedures and rigorous analysis
- emphasizes the development of generalizations, principles or theories and is reported to others interested in the problem
- refines the problem or questions as the research progresses

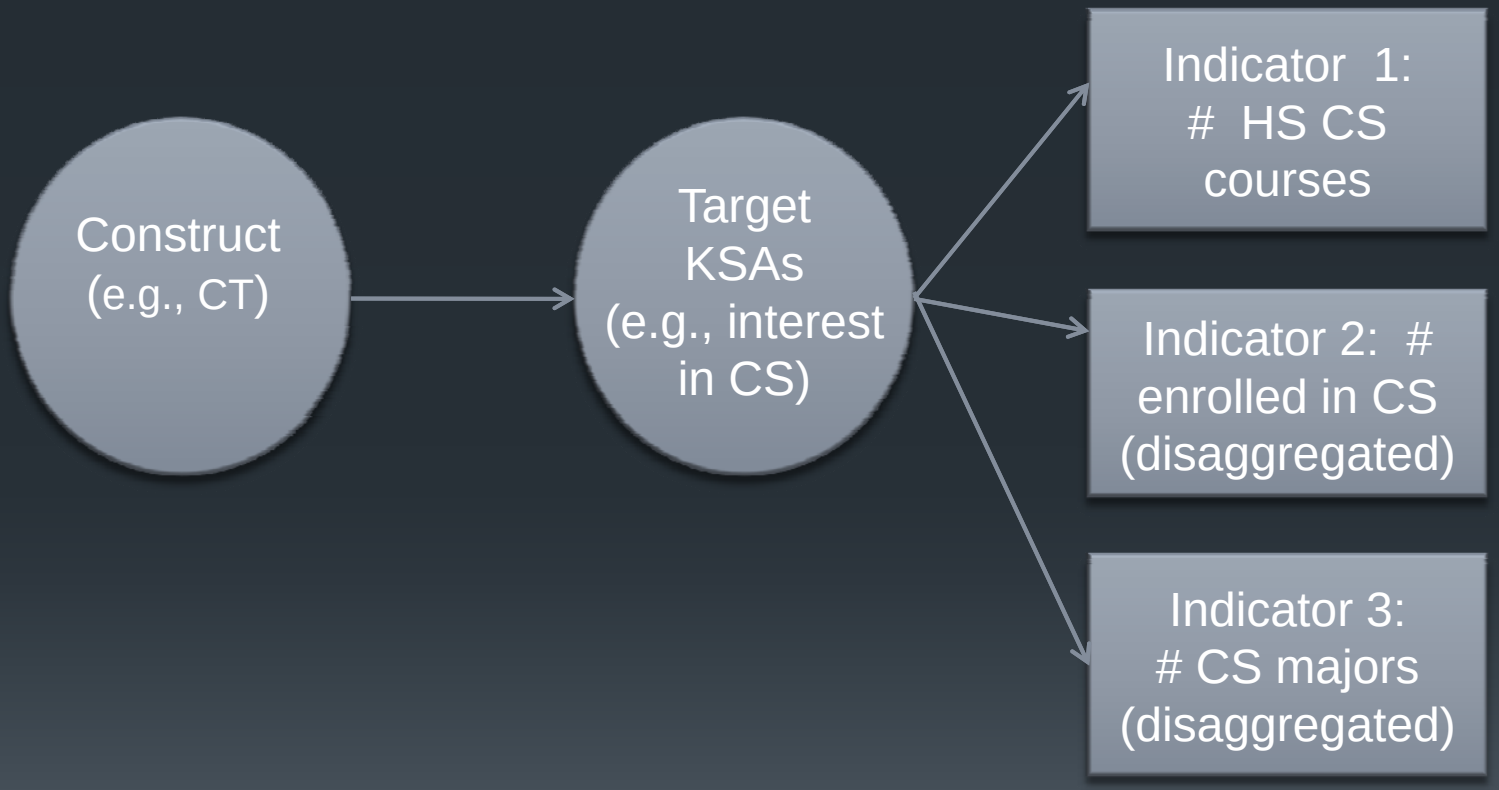
*Retrieved from http://www.nsf.gov/pubs/2011/nsf11034/nsf11034.jsp

To Strengthen the Coherence of Your Proposal…

- ✓Make explicit connections among:
  1. Intellectual merit
  2. Broader Impact
  3. Key program activities
  4. Plan for evaluation
- ✓Leverage use of graphic organizers
  1. To aid reviewers
  2. As a foundation for the evaluation plan

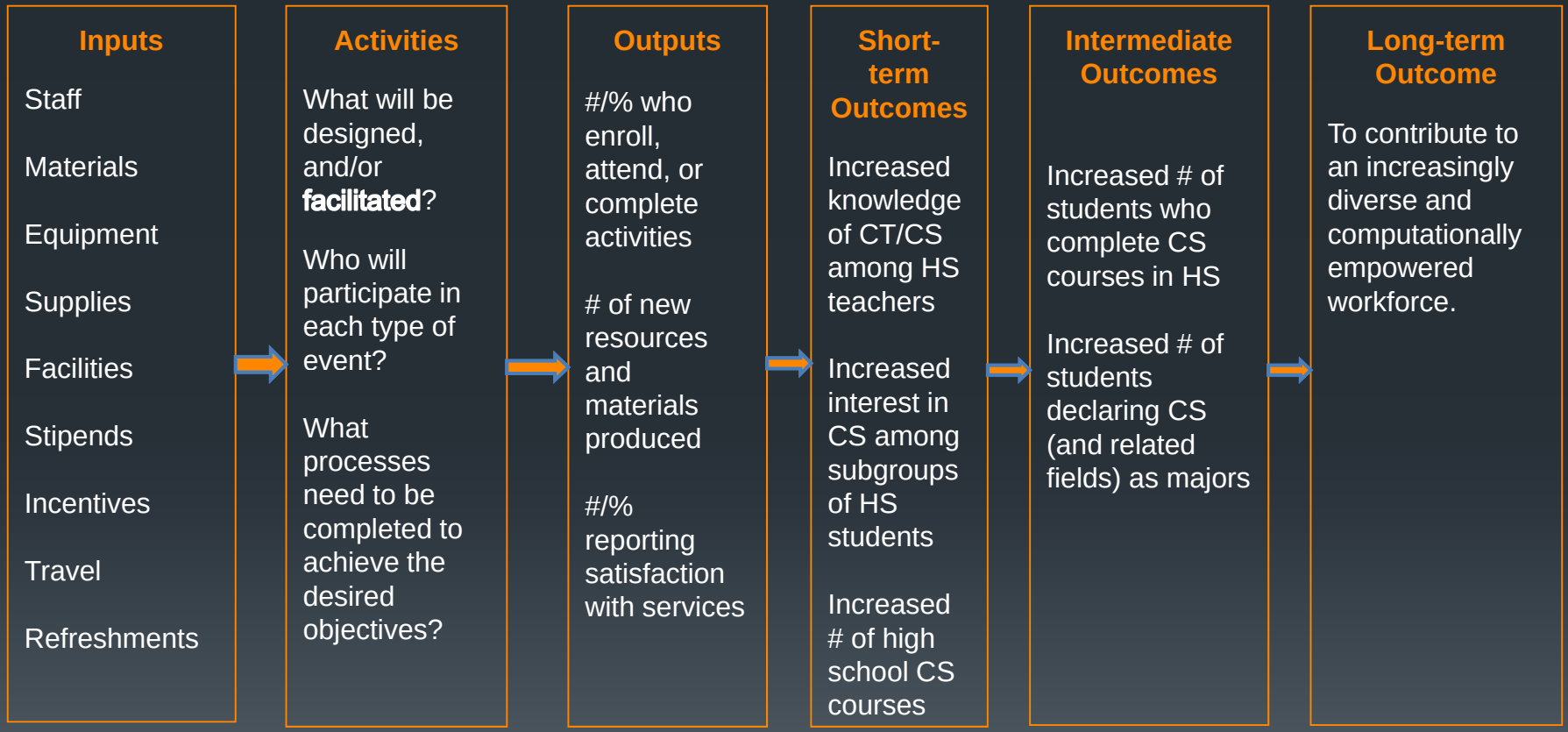
Using a Theory of Action to Summarize Your Theory about Existing Theories…

Building on Your Theory of Action to Support Evaluation



Using a Program Logic Model

***CE21 Program Goal**: contribute to understanding how diverse student populations are engaged and retained in computing, learn its fundamental concepts, and develop computational competencies.



*From the CE21 Program Synopsis http://www.nsf.gov/pubs/2012/nsf12527/nsf12527.pdf

The Role of Assessment in a Summative Evaluation

Evaluation Question	Indicator or Benchmark	Data Sources	Methods
1. Did the project produce targeted changes in *knowledge and attitudes**?	#/% who demonstrate knowledge CT gains and/or more positive attitudes or beliefs towards CS	•**Content assessments** •Attitude surveys •Interviews & focus groups	Quantitative & Qualitative
2. Did project participation lead to desired changes in *behaviors**?	#/% who apply CT concepts to solve real-world problems	•**Ability-based assessments** •Observations •Journal entries •Surveys	Quantitative & Qualitative
3. Did the project result in enhancements to infrastructure?	#/type of facilities #/type of instrumentation #/type of networks and partnerships	•Site visits •Document review •Dissemination products (conferences, publications, web materials)	Quantitative & Qualitative

*How do outcomes compare, when disaggregated by subgroup (e.g., gender, ethnicity, SES, education)?

Practical Considerations When Selecting Assessments

- ✓Are existing assessments aligned with key constructs & outcome indicators?
  - ✓Preserving existing psychometric properties
  - ✓“Customizing” use of existing assessments
- ✓Are resources adequate to develop a more sensitively-aligned assessment?
  - ✓What else needs to be considered?

Tying it Together: Aligning Proposed Evaluations with Efforts to Map CT Domain

- Use a Theory of Action to
  - Build on (or challenge) existing studies
  - Integrate research about STEM education
  - Identify how your proposal will move knowledge forward

- Specify Measurable Indicators to Document Potential Impact
  - Analyze outcomes disaggregated by subgroup
  - Develop a plan for dissemination
  - Leverage ways to sustain key components after funding ends

- Use Graphic Organizers to Depict
  - Links between the proposed program and research (theory of action)
  - How funds will be used and what will be accomplished (logic model)
  - What will serve as evidence of progress towards goals (evaluation plan)

Questions & Comments

Now it's your turn. Please feel free to ask questions and comment on your work in this area.

Also consider:

- In what ways have you successfully and unsuccessfully used assessment in your CT program evaluations?

Activity 3
Tasks/Situations for Eliciting Evidence of Focal CT Proficiencies

We want to hear from people in the field:

> *What **types of tasks or situations** would elicit evidence of important CT proficiencies?*

Please review the Activity 3 worksheet.

Work in small groups, or individually, to identify possible types of tasks or situations that would elicit evidence for a subset of CSTA standards.

We encourage you to consider and discuss the contexts in which you work as you identify possible tasks or situations that would elicit the desired evidence.

Closing Comments

- Computational Thinking is an emerging domain whose boundaries can and should vary across contexts.
- Valid assessment of CT proficiencies requires carefully specifying and aligning:
  - ✓ CT knowledge, skills, attitudes
  - ✓ Evidence of the knowledge, skills attitudes
  - ✓ Tasks or situations for eliciting the desired evidence
- ✓ Assessments **are not** plug n play, **are** use specific, and should focus on evidence, not just tasks.
- ✓ Evaluation often uses assessments to measure targeted knowledge, skills, and attitudes, and is a critical component of competitive NSF proposals.

THANK YOU!