Games, Learning, and Stealth Assessment

Valerie Shute

Webinar, April 25, 2012

Our Team!







Games, Learning, Assessment

Claim 1

Good games can act as *transformative environments* to support skill development and deep, meaningful learning.

Claim 2

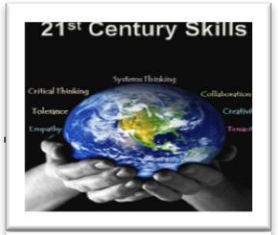
Learning is best when it is active, interesting, goal-oriented, and contextualized (i.e., features of good games).

Claim 3

Stealth assessment can collect dynamic evidence of learning in real-time, at various grain sizes (and use info to support learning).

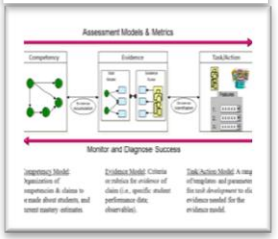
My Game Plan

- Talk about 21st century competencies……



- Discuss games and learning……

- Overview evidence-centered design……



- Describe stealth assessment……

- Illustrate approach in current project……


21st Century
Competencies

It’s Time for a Change

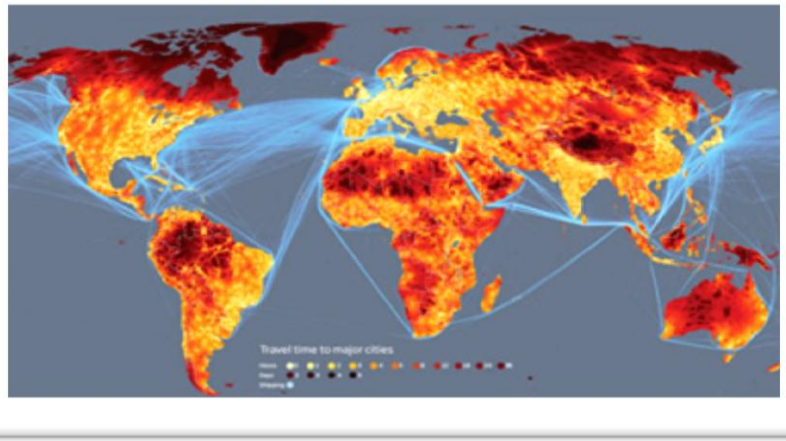
The world has changed a lot in the past 100 years. Education has not.

Classroom photo, 1910

Classroom photo, 2010

21st Century Skills

- Today's kids need new skills to be successful and productive (e.g., creativity, collaboration, communication).
- The world in which we live is complex and interconnected.




Collaboration
Critical Thinking
Systems Thinking
Persistence
Creativity
Curiosity
Empathy


Collaboration
Critical Thinking
Systems Thinking
Persistence
Creativity
Curiosity
Empathy

Challenges

- One problem with embracing new skills is the lack of valid & reliable assessments for them.
- Old ways of testing won't work. Learning and succeeding in our complex world can't be optimally be measured by MC tests or self-report.
- *Re-think assessment!*


Games

Faces of engagement…

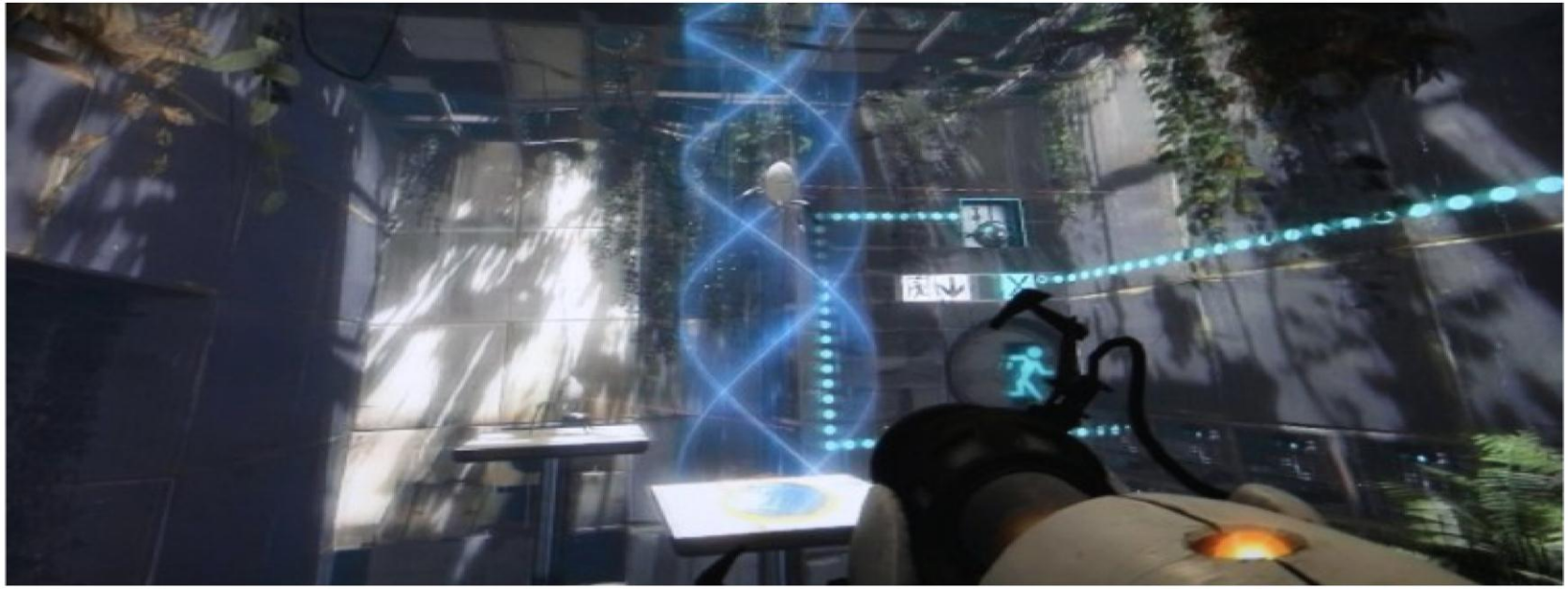
… on kids of all ages!

Engagement here?

Not so much.

Why games as assessments?

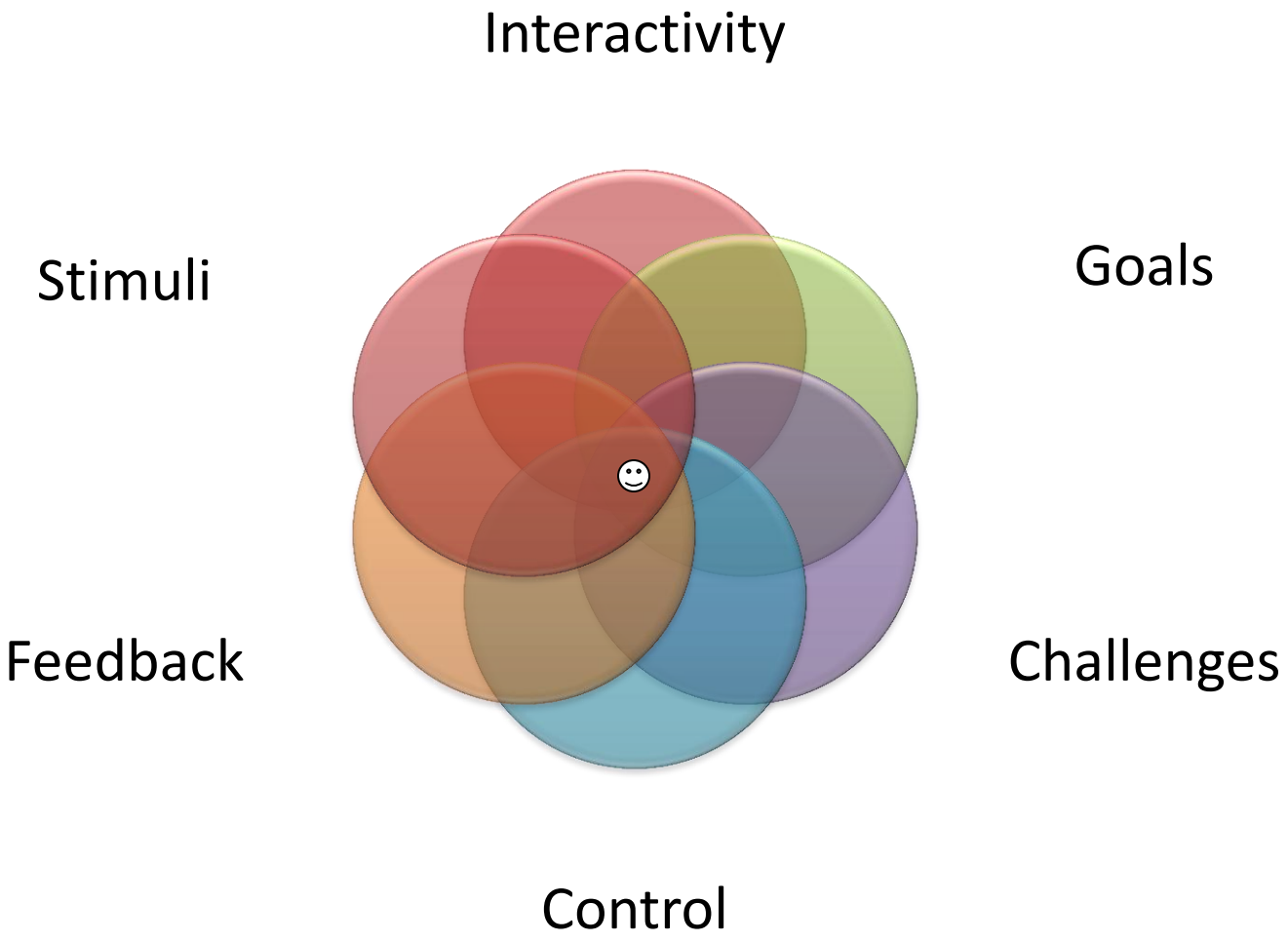
1. Good games are *engaging* and require a player to *apply various competencies* to succeed (e.g., problem solving).
2. Games are also *ubiquitous*. A survey of 1,102 youth (ages 12 - 17) found that 97% play video games: boys=99%; girls=94% (Lenhart et al., 2008).



Good Game Elements

1. Interactive problem solving
2. Specific goals/rules
3. Adaptive challenges
4. Control
5. Ongoing feedback
6. Sensory Stimuli

Gestalt of Games




Learning

Learning

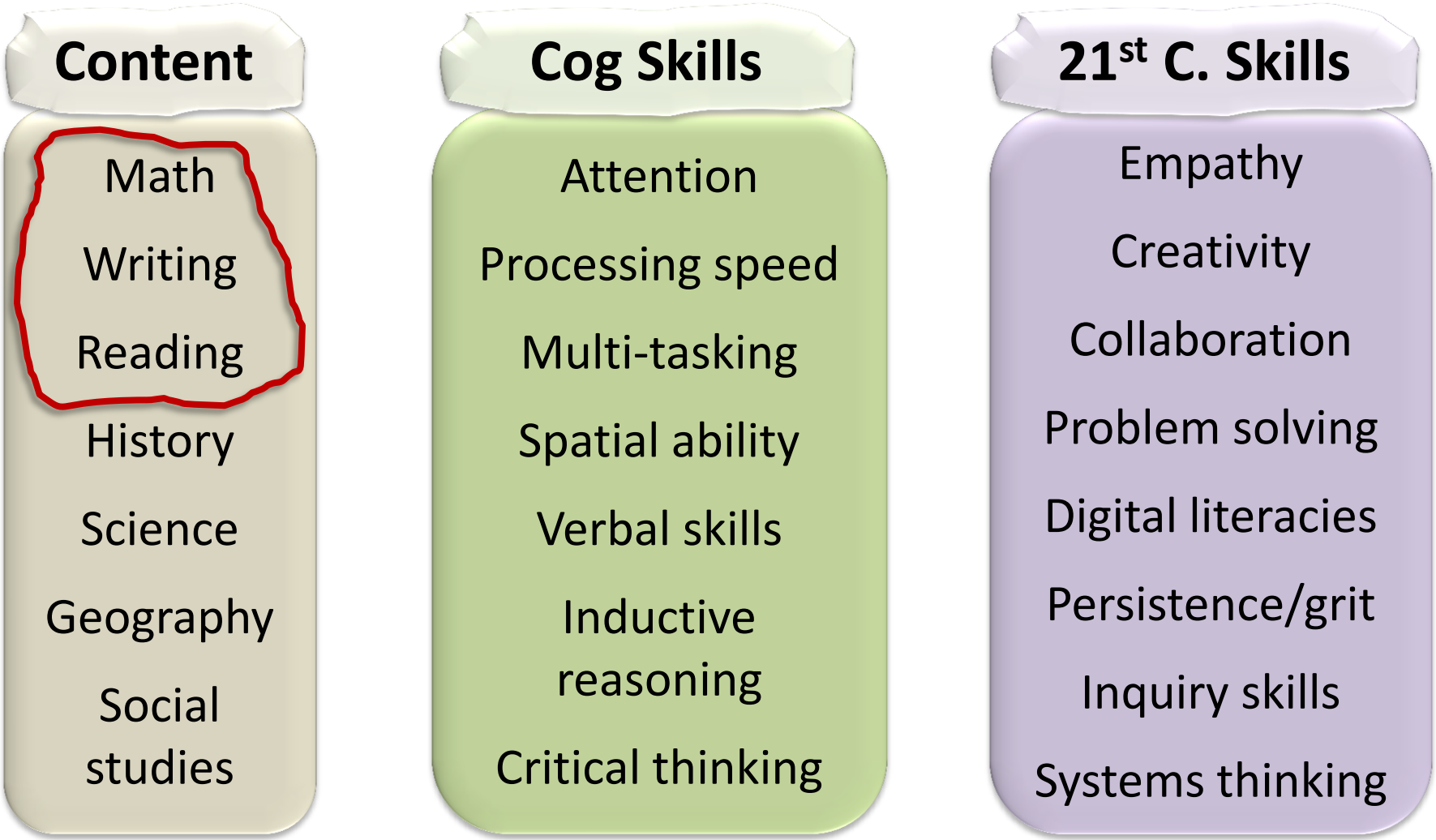
- ***Lifelong process*** of accessing, interpreting, and evaluating info & experiences into knowledge, skills, values, dispositions, etc.

- ***Change*** from one point in time to another in terms of knowing, doing, believing, feeling, etc.

Learning Theories

- Constructivism and situated learning.
- Learner is active in the learning process; learning is the result of interaction with a problem context where learners construct meaning.

Kinds of Learning



Learning Outcomes

20th C. Outcomes

↑ Test scores

↑ Test scores

↑ Test scores

↑ Test scores

↑ Test scores

↑ Test scores

↑ Test scores

21st Century Outcomes

↑ College enrollment

↑ Digital literacies

↑ Kindness

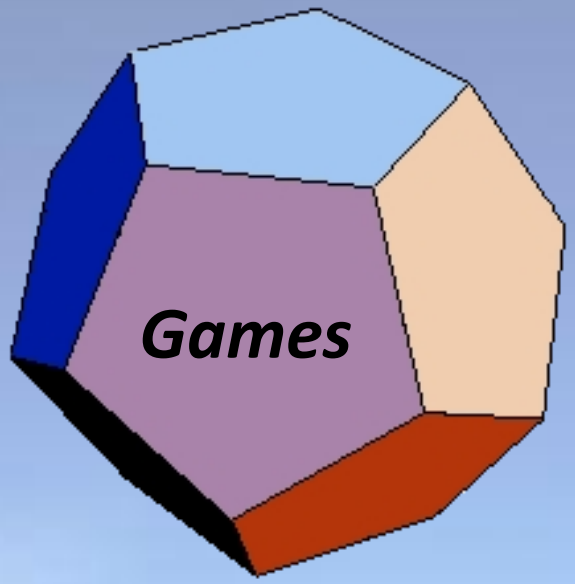
↓ HS dropouts

↓ Intolerance/Bigotry

↑ Adaptivity

↑ Civic engagement

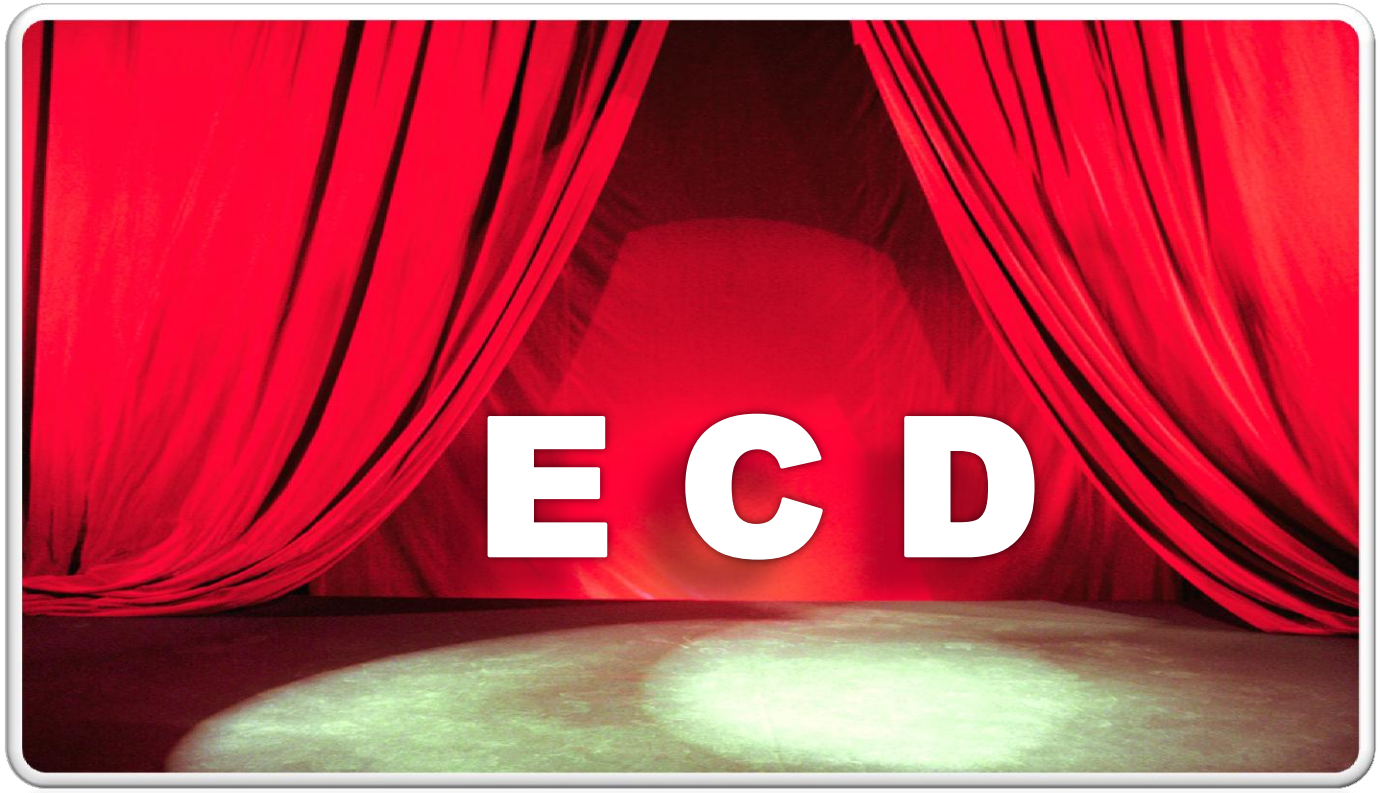
↑ Happiness



Narrow View of Learning


Assessment
Design

Evidence-Centered Design



(e.g., Mislevy, Steinberg, & Almond, 2003)

Assessment Design

Competency Model

What do you want to say about the person?

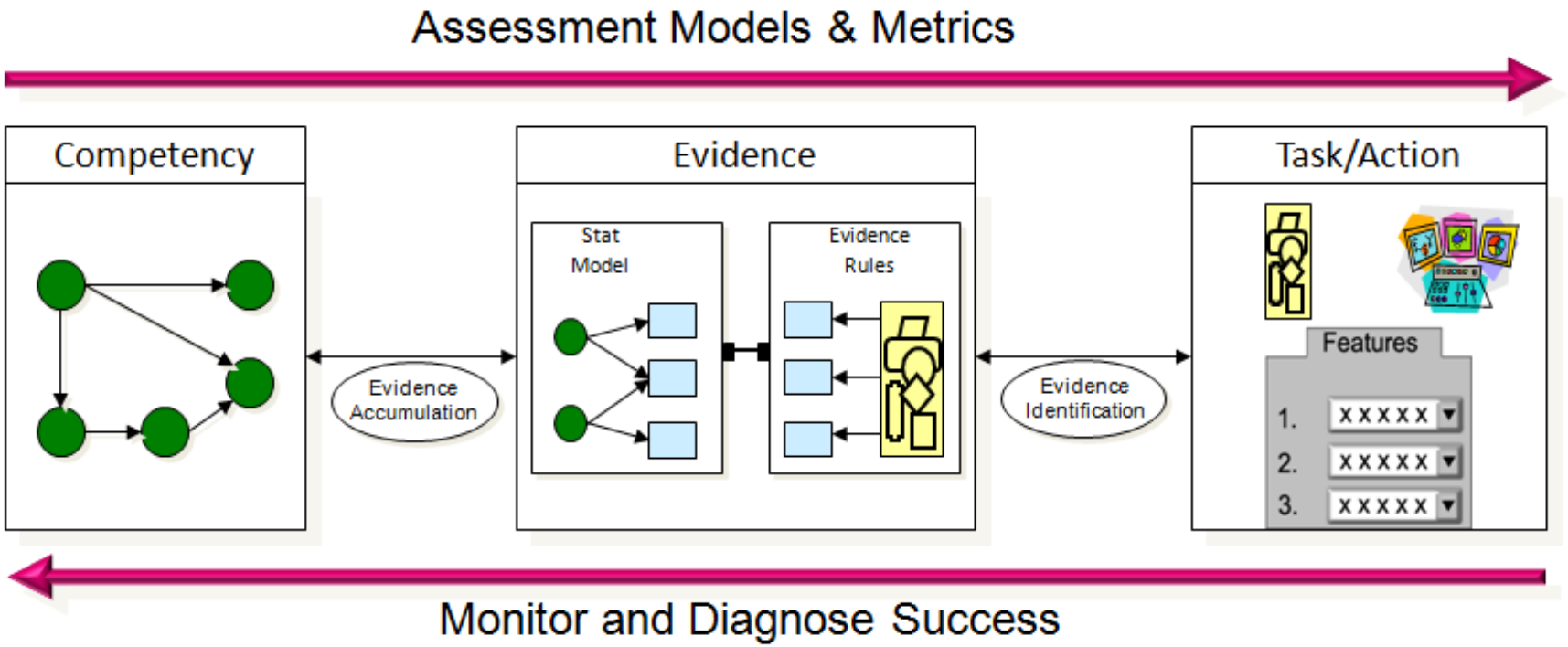
Evidence Model

What observations would provide best evidence for what you want to say?

Task/Action Model

What kinds of tasks let you make the necessary observations?

Design & Diagnosis



Competency Model: Organization of competencies & claims to be made about students, and current mastery estimates.

Evidence Model: Criteria or rubrics for *evidence* of claim (i.e., specific student performance data; observables).

Task/Action Model: A range of templates and parameters for *task development* to elicit evidence needed for the evidence model.

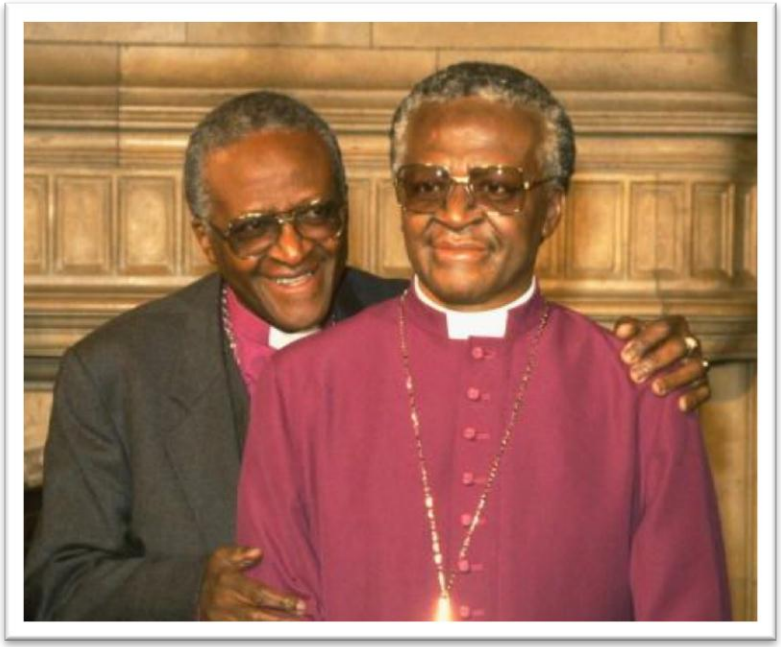

Stealth
Assessment

Stealth Assessment

- **Advances in measurement let us administer evidence-based assessments to:**
  - **Extract *ongoing* information from a learner**
  - **Make *accurate inferences* of competencies**
  - **React in *immediate* and helpful ways.**
- **Accomplished via automated scoring and machine-based reasoning techniques.**
- **When assessment is so seamlessly woven into the fabric of the learning environment that it's invisible, this is *stealth assessment*.**

Stealth Assessment Features

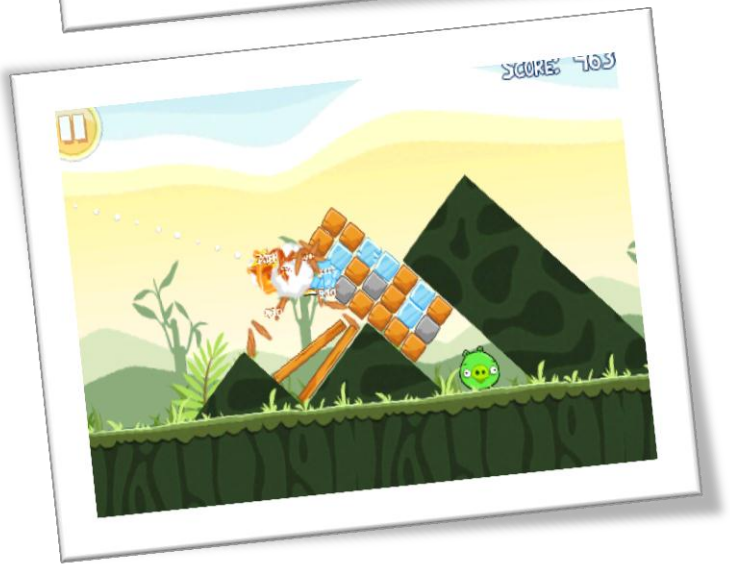
Seamless & Ubiquitous Assessment



Accurate & Rich Learner Models

Stealth Assessment Process

- During gameplay, players produce rich sequences of actions while performing complex tasks.
- These actions draw on a variety of competencies.
- Evidence needed to assess competencies is provided by the players' interactions with the game itself (i.e., the processes of play).



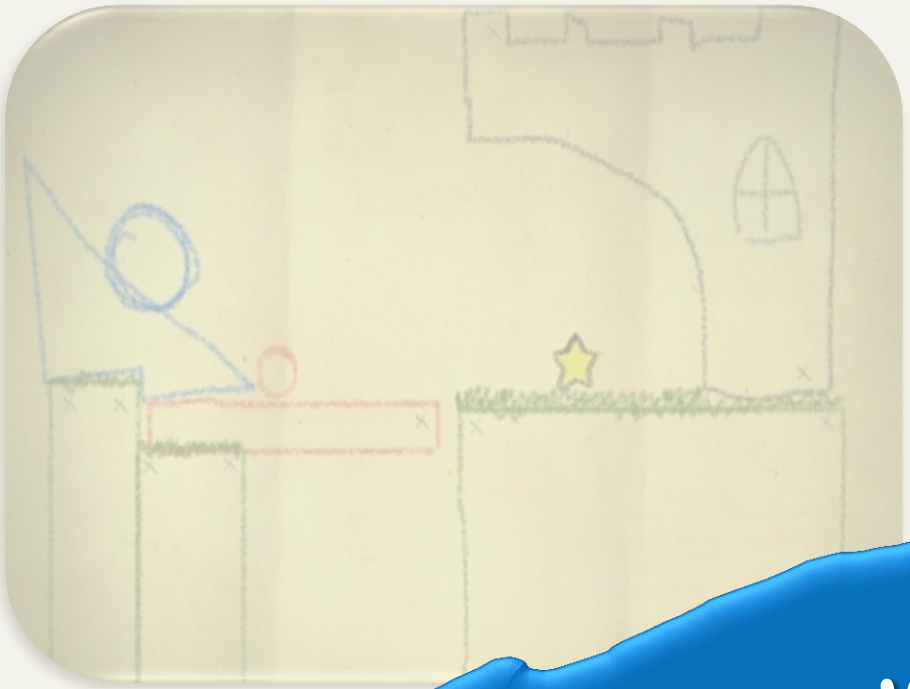
Stealth Assessment Process

Stealth assessment allows us to:

- *Monitor a player's current level* on valued competencies (aggregating data across multiple instances)
- *Provide immediate feedback* (explicit, like status bars; implicit, like dying) to player/teacher/etc.
- *Use info as basis for support*, e.g., adjusting difficulty level of challenges, providing status report for the teacher and/or student.

Example

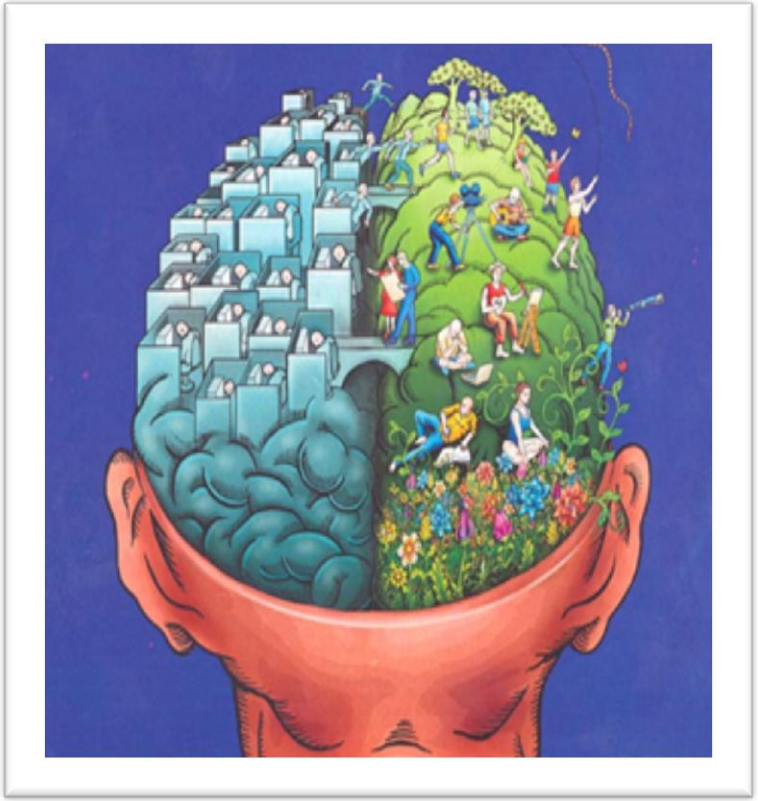
BILL & MELINDA GATES foundation




Crayon Physics Deluxe

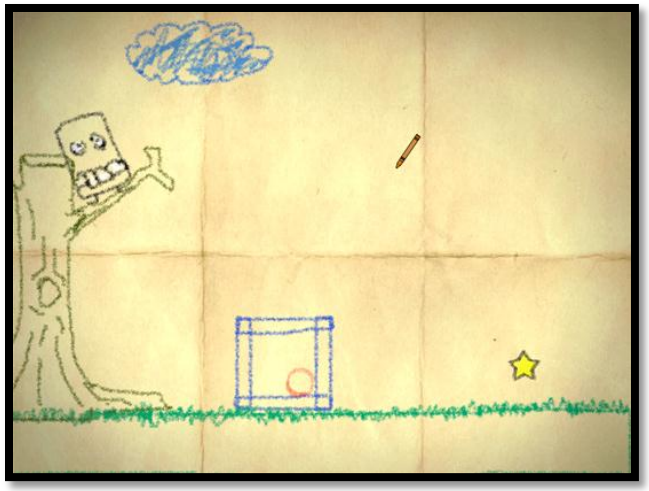
Project Overview

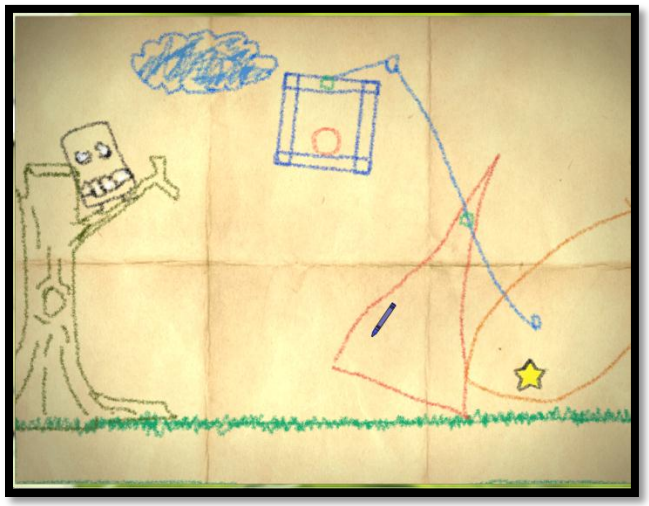
- Develop and evaluate assessments for *creativity*, *persistence*, & *conceptual physics*.
- Embed assessments in Crayon Physics Deluxe (CPD), a game with 2D physics simulations.
- Data will be collected in CPD from players' interactions with the game. These data serve as the basis for the 3 “stealth assessments.”



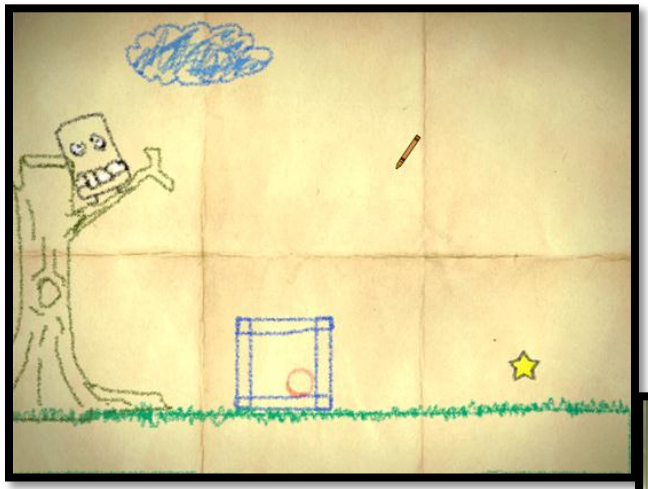
Crayon Physics Deluxe

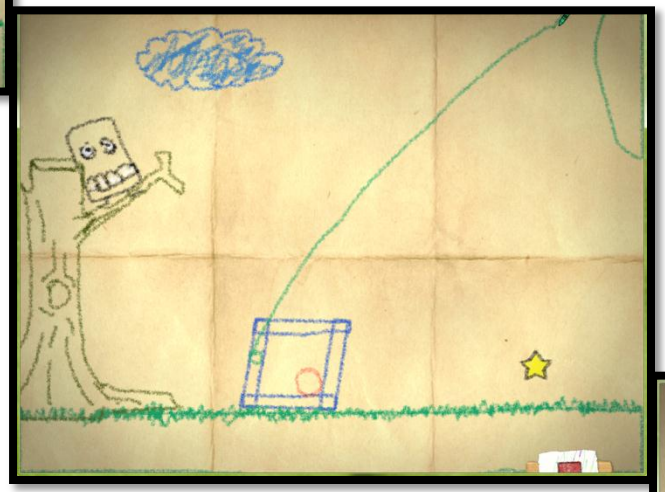
- Objective: guide **red ball** from a starting point to a star . Everything obeys the basic rules of physics (e.g., gravity and Newton's three laws of motion).
- Player draws physical objects that "come to life" when drawn (e.g., levers, ramps, pendulums) in order to get the ball to the star.
- CPD lets players solve problems in many different ways. *"It is not about finding just any solution. It is about finding the **awesomest** one."*

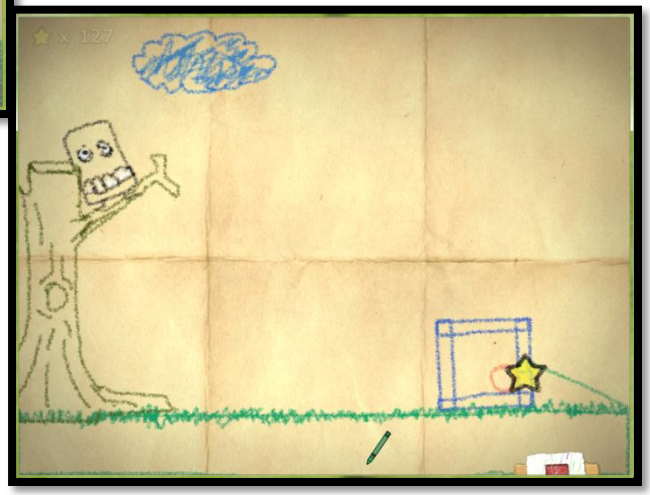




Or….

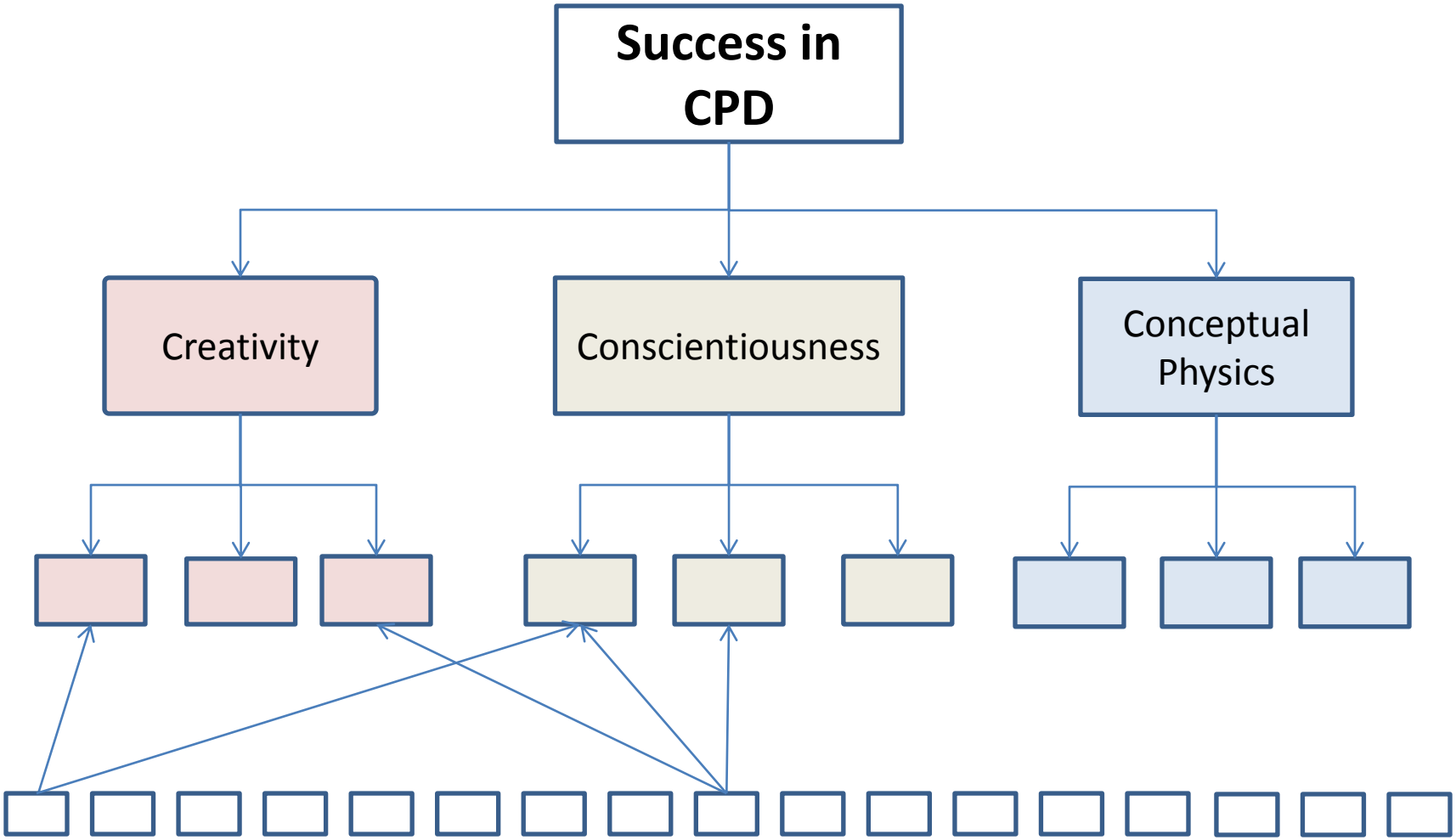






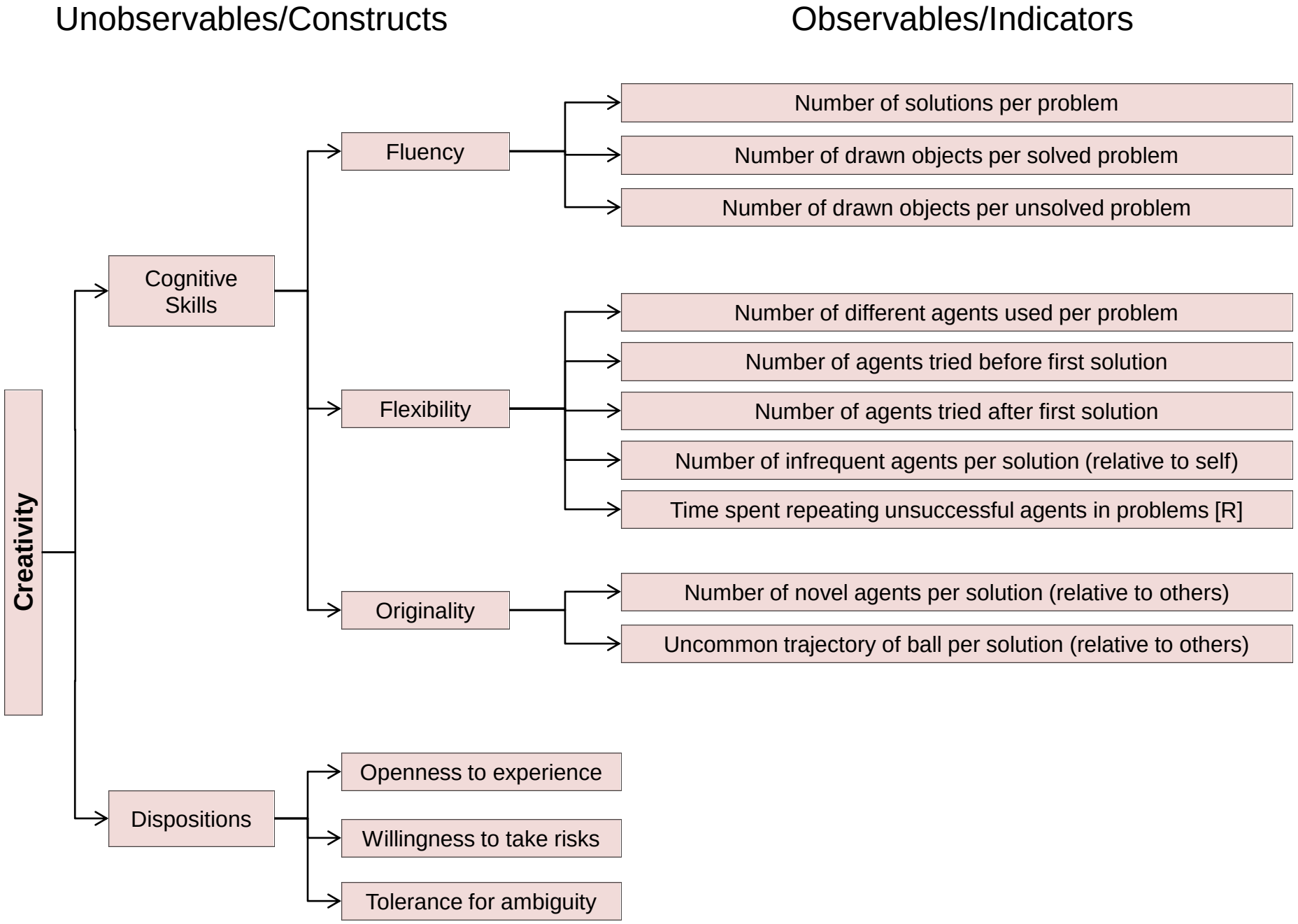
Research Questions

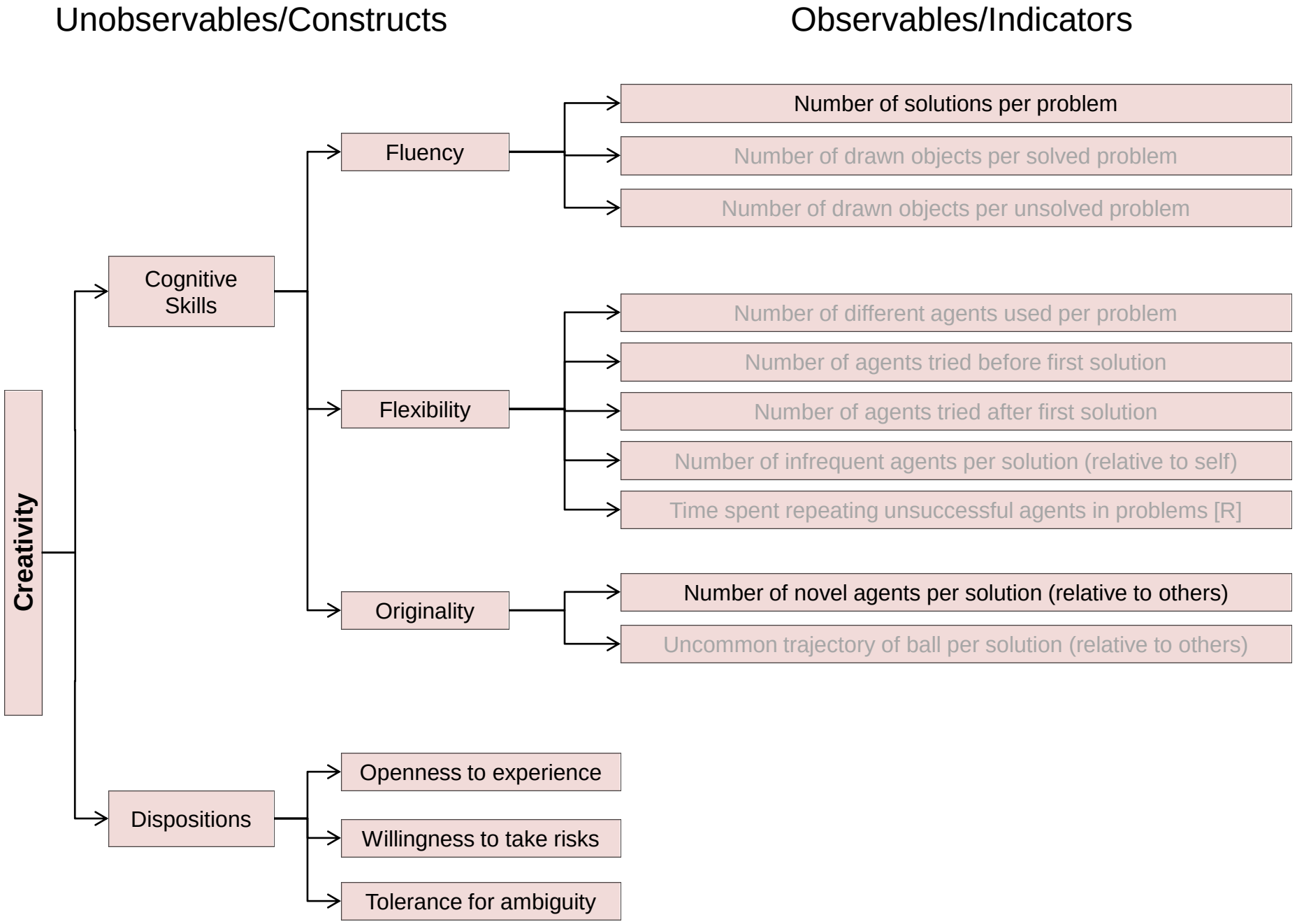
- ***Valid/Reliable?*** Are the intrinsic assessments we develop in *Crayon Physics Deluxe* valid & reliable measures of our 3 focal competencies?
- ***Learning?*** Do students learn conceptual physics from playing the game?
- ***Portable?*** Are the ECD-based models we develop for Crayon Physics transferable to another context/environment/game?


Success in CPD
Creativity
Conscientiousness
Conceptual Physics

Creativity

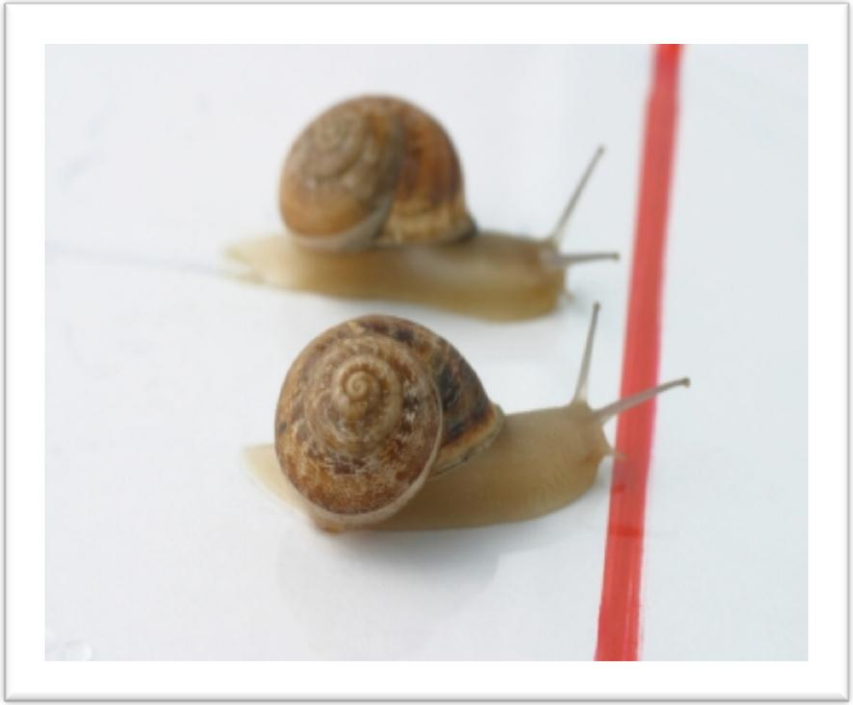
- Creativity is vital for innovation (e.g., Weisberg, 2004).
- Needed in science, technology, humanities, and the arts, as well as in organizations and businesses.
- We define creativity as having cognitive (e.g., novelty, fluency) and dispositional (e.g., openness, risk-taking) aspects.

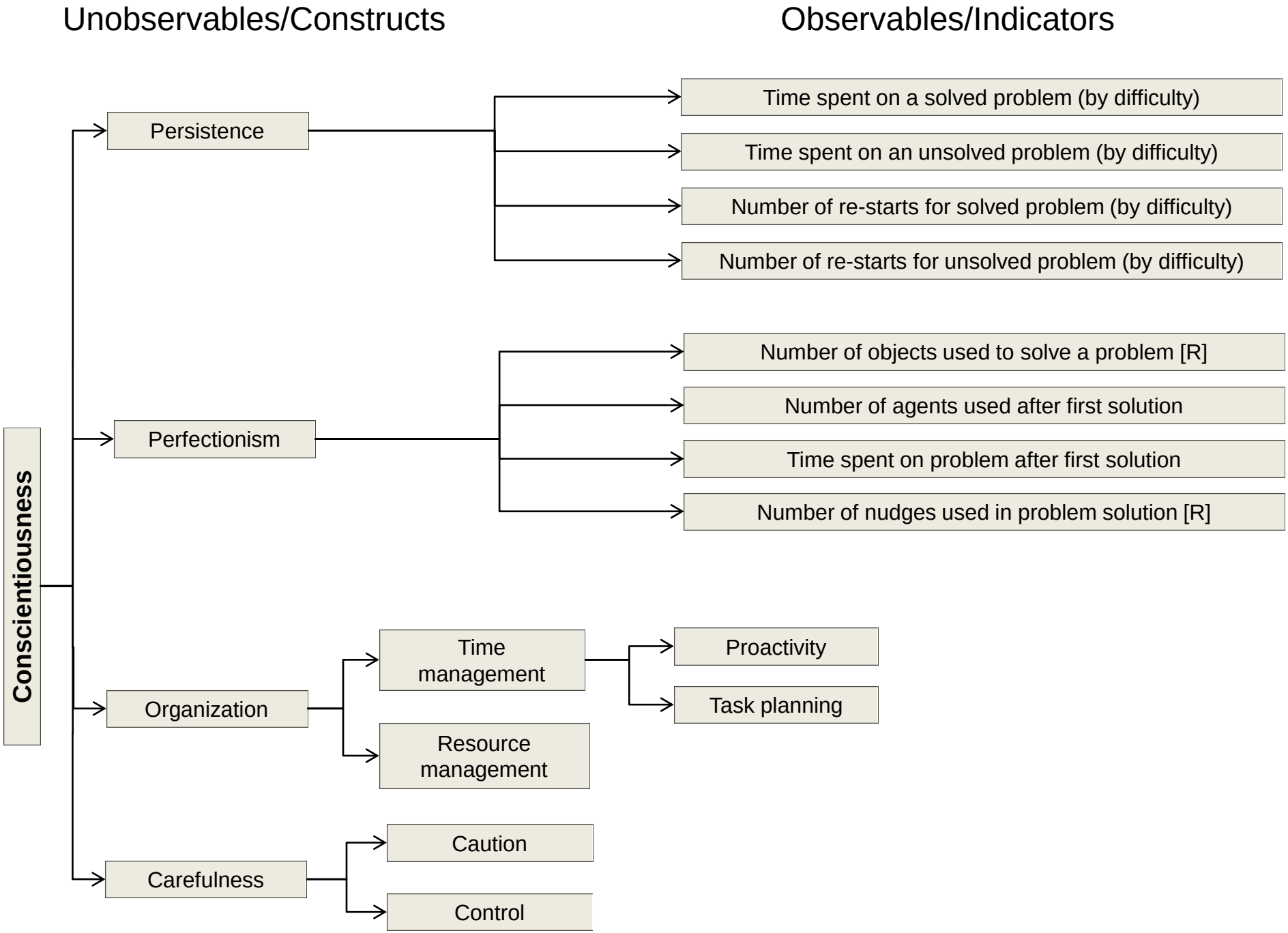

Unobservables/Constructs
Observables/Indicators
Creativity
Cognitive Skills
Dispositions
Fluency
Flexibility
Originality
Openness to experience
Willingness to take risks
Tolerance for ambiguity
Number of solutions per problem
Number of drawn objects per solved problem
Number of drawn objects per unsolved problem
Number of different agents used per problem
Number of agents tried before first solution
Number of agents tried after first solution
Number of infrequent agents per solution (relative to self)
Time spent repeating unsuccessful agents in problems [R]
Number of novel agents per solution (relative to others)
Uncommon trajectory of ball per solution (relative to others)

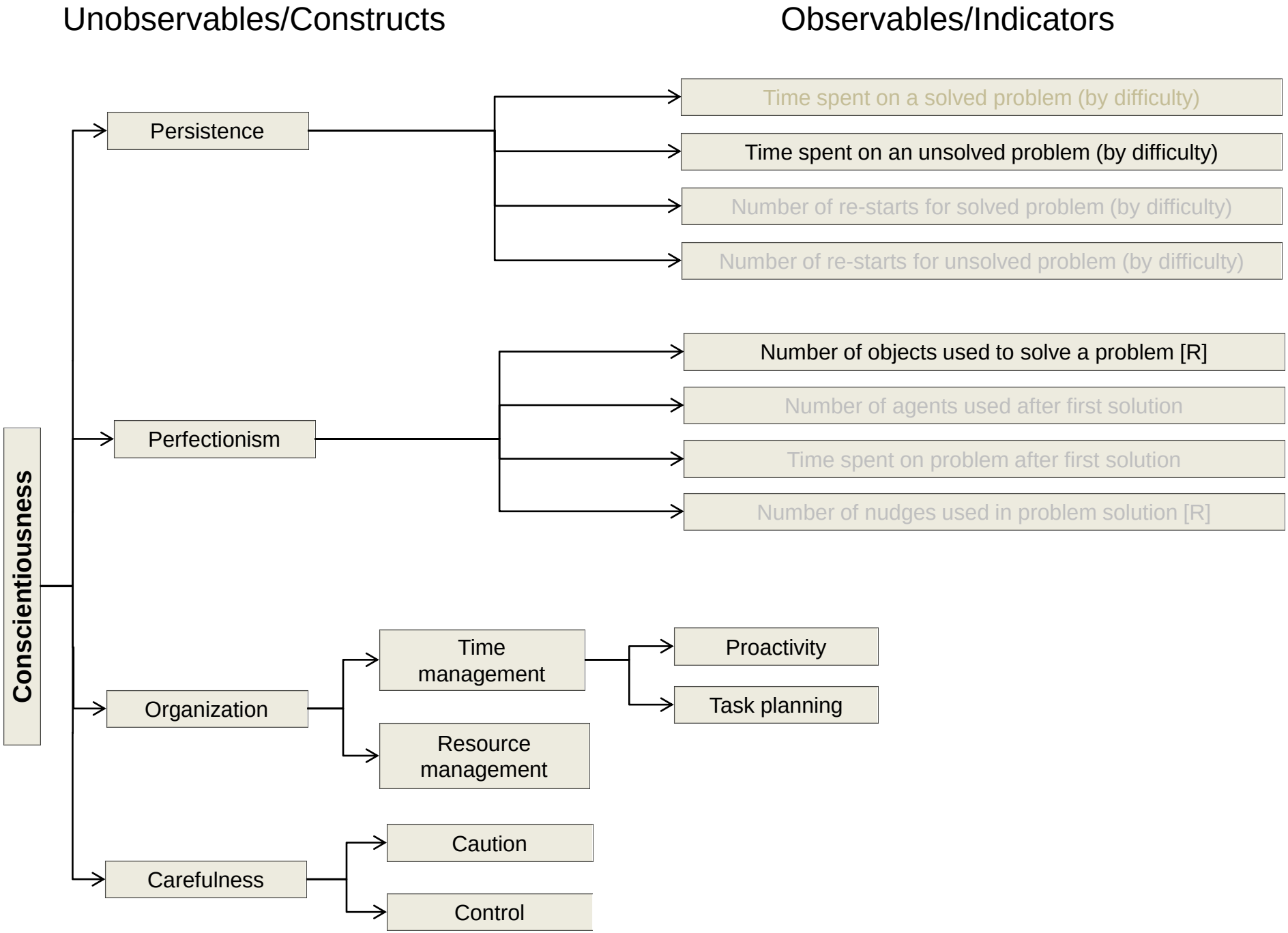

Unobservables/Constructs
Observables/Indicators
Creativity
Cognitive Skills
Dispositions
Fluency
Flexibility
Originality
Openness to experience
Willingness to take risks
Tolerance for ambiguity
Number of solutions per problem
Number of drawn objects per solved problem
Number of drawn objects per unsolved problem
Number of different agents used per problem
Number of agents tried before first solution
Number of agents tried after first solution
Number of infrequent agents per solution (relative to self)
Time spent repeating unsuccessful agents in problems [R]
Number of novel agents per solution (relative to others)
Uncommon trajectory of ball per solution (relative to others)

Conscientiousness

- Predicts academic achievement from preschool to high school, to the postsecondary level, and adulthood.
- We define it as the *motivation to work hard despite challenging conditions*; and being organized, thorough, and careful.

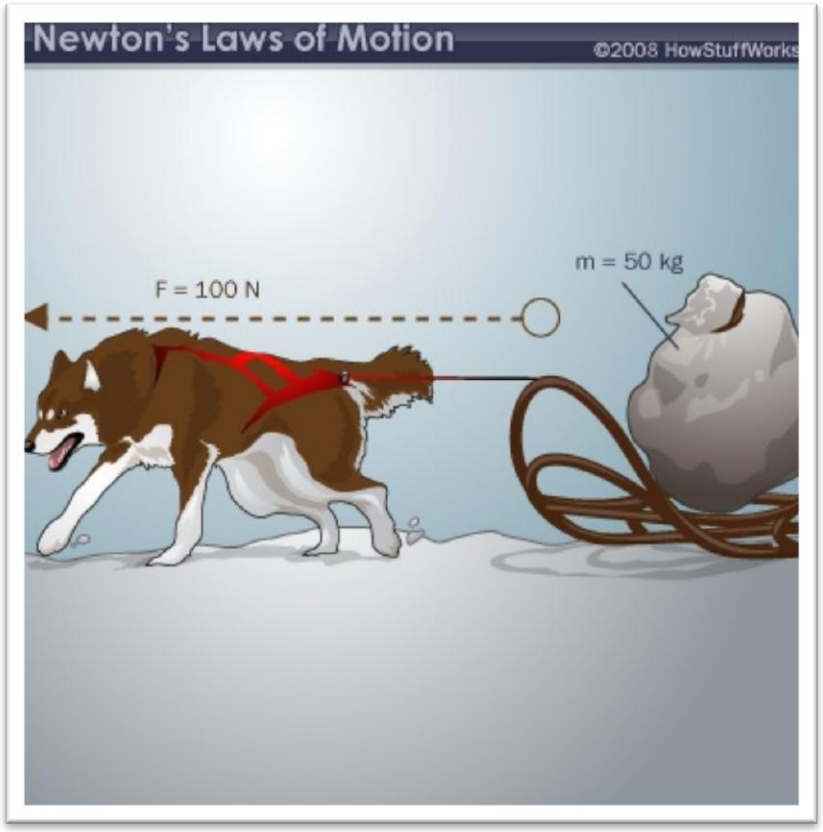



Unobservables/Constructs
Observables/Indicators
Conscientiousness
Persistence
Perfectionism
Organization
Carefulness
Time management
Resource management
Caution
Control
Proactivity
Task planning
Time spent on a solved problem (by difficulty)
Time spent on an unsolved problem (by difficulty)
Number of re-starts for solved problem (by difficulty)
Number of re-starts for unsolved problem (by difficulty)
Number of objects used to solve a problem [R]
Number of agents used after first solution
Time spent on problem after first solution
Number of nudges used in problem solution [R]


Unobservables/Constructs
Observables/Indicators
Conscientiousness
Persistence
Perfectionism
Organization
Carefulness
Time spent on a solved problem (by difficulty)
Time spent on an unsolved problem (by difficulty)
Number of re-starts for solved problem (by difficulty)
Number of re-starts for unsolved problem (by difficulty)
Number of objects used to solve a problem [R]
Number of agents used after first solution
Time spent on problem after first solution
Number of nudges used in problem solution [R]
Time management
Resource management
Proactivity
Task planning
Caution
Control

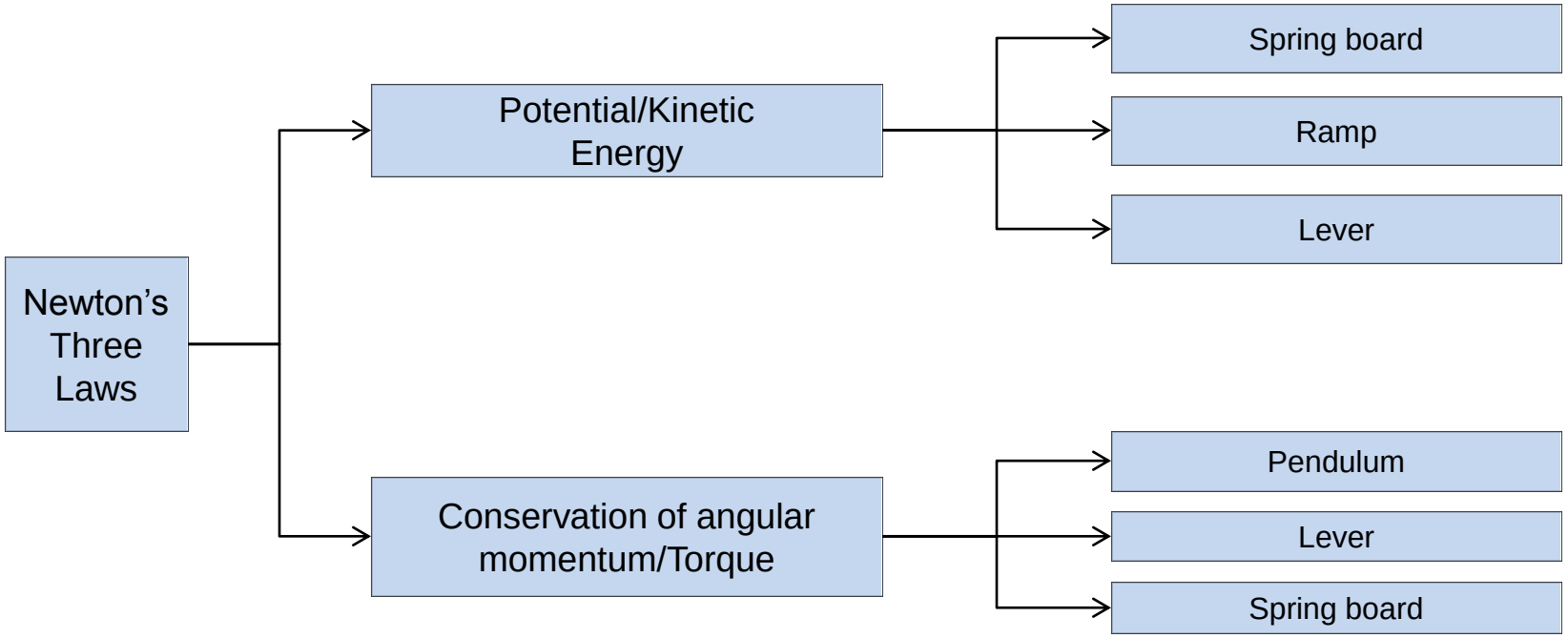
Conceptual Physics

- Understanding physics concepts is key to understanding the physical world.
- Misconceptions abound. For example, Halloun and Hestenes (1985) found that *only 15%* of their 478 college physics students showed an accurate understanding of the relationship between unbalanced forces and acceleration (Newton's 2nd law: F = ma).



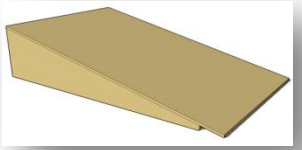
Unobservables/Constructs

Agents of Force and Motion

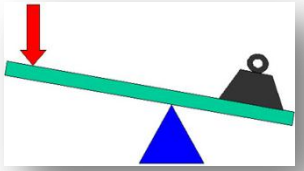


Agents of Force/Motion

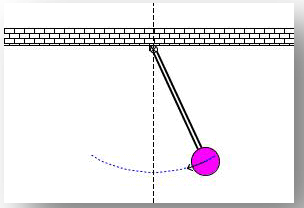
- ***Ramp***: Used to change the direction of the motion of the ball (or another object).



- ***Lever***: Rotates around a fixed point usually called a fulcrum or pivot point.



- ***Pendulum***: Directs an impulse tangent to its direction of motion.



- ***Springboard***: Stores elastic potential energy provided by a falling weight. Elastic potential energy becomes kinetic as the weight is released.

Micro-indicators for Agents

Agents	Micro-Indicators
Ramp	1. Number of bends (or tubes, i.e., tortuosity) 2. Angle of each bend 3. Length of ramp
Lever	1. Length of the lever 2. Position of fulcrum 3. Height through which object falls before hitting lever 4. Mass of object 5. Location of the dropped object on lever (distance from fulcrum)
Pendulum	1. Angle of pendulum relative to horizontal fulcrum (180 degrees is max) 2. Length between the axis point and the fulcrum (Moment of inertia) 3. Mass (important when the pendulum hits something) 4. Position of pin
Springboard	1. Length of springboard 2. Mass of the object to weight it down 3. Position of the ball at release 4. Delete object or let fall off springboard 5. Angle of springboard at release (90 degrees max)

Micro-indicators for Agents

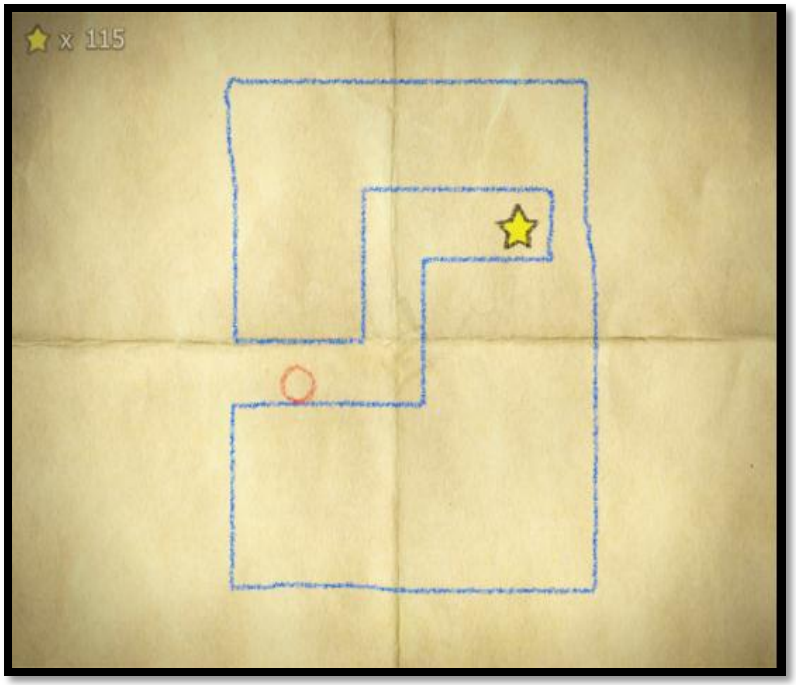
Agents	Micro-Indicators
Ramp	1. Number of bends (or tubes, i.e., tortuosity) 2. Angle of each bend 3. Length of ramp
Lever	**1. Length of the lever** **2. Position of fulcrum** **3. Height through which object falls before hitting lever** **4. Mass of object** **5. Location of the dropped object on lever (distance from fulcrum)**
Pendulum	1. Angle of pendulum relative to horizontal fulcrum (180 degrees is max) 2. Length between the axis point and the fulcrum (Moment of inertia) 3. Mass (important when the pendulum hits something) 4. Position of pin
Springboard	1. Length of springboard 2. Mass of the object to weight it down 3. Position of the ball at release 4. Delete object or let fall off springboard 5. Angle of springboard at release (90 degrees max)

CPD Difficulty Indices

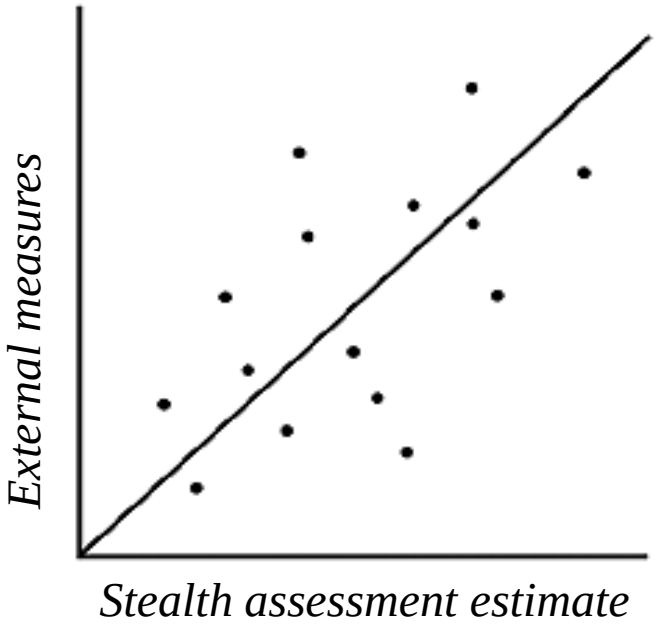
- ***Relative location of ball to star***. If star is above the ball in a problem, this forces the player to use a lever, pulley, springboard, or pendulum to solve the problem (0-1 point).
- ***Obstacles***. If the pathway between ball and star is obstructed, this requires the player to project the ball in a very specific trajectory to obtain the star (0-2 points).
- ***Distinct agents of force/motion***. A CPD problem may require one or two agents to get ball to the star (0-1 point).
- ***Novelty***. This addresses whether a problem is novel relative to other problems played. Problem solution is not easily determined from experience with other problems (0-2 points).

Difficulty Example

- Each problem evaluated under all of the rubrics to yield a total difficulty score (i.e., from 0-6).
- Cave story problem is below. Its difficulty score = ***5*** (hard!) as the star is above the ball (1), there's one obstacle which is a narrow pathway (1), two agents are typically needed to solve it (1) and there's no other problem like that in the game (2).

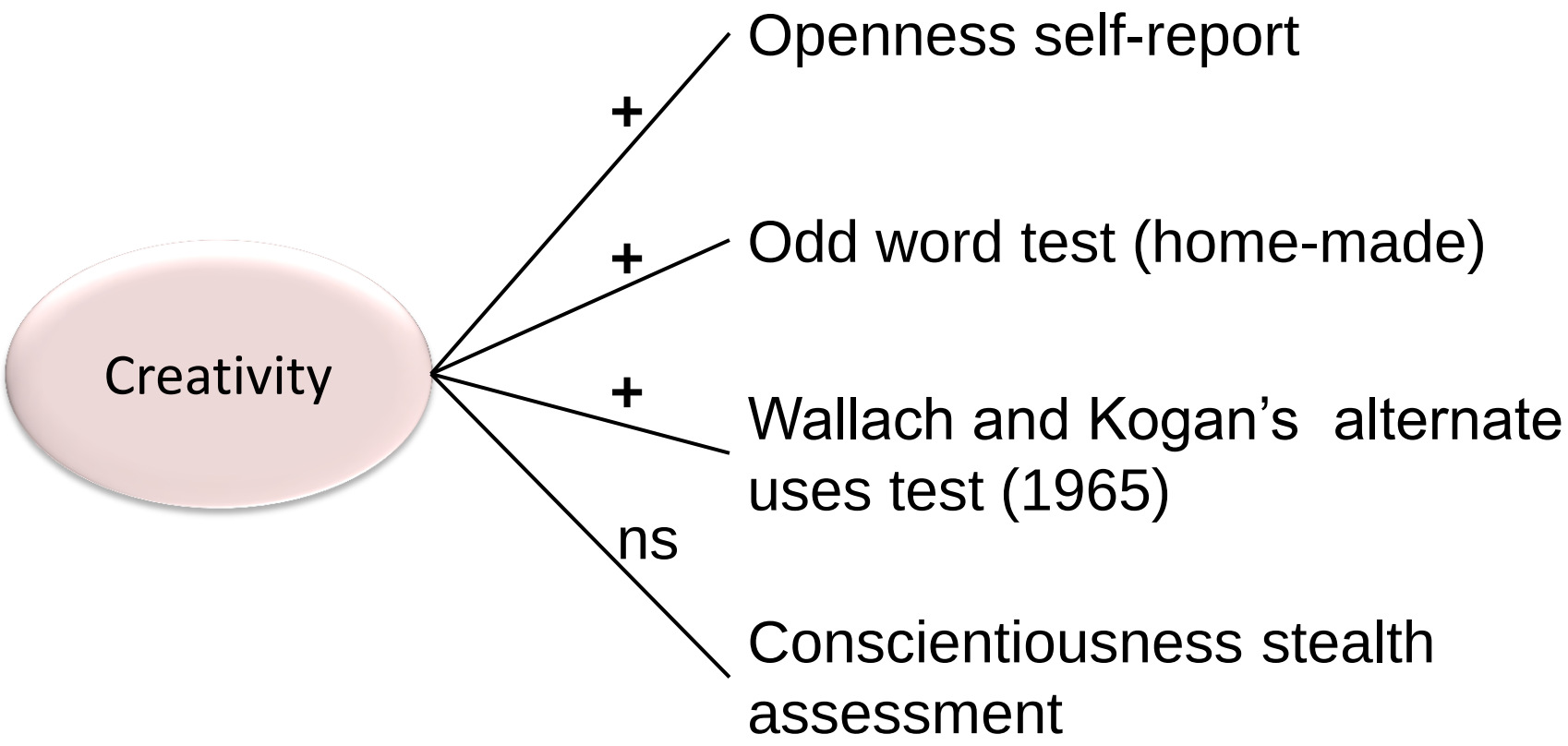


Validation



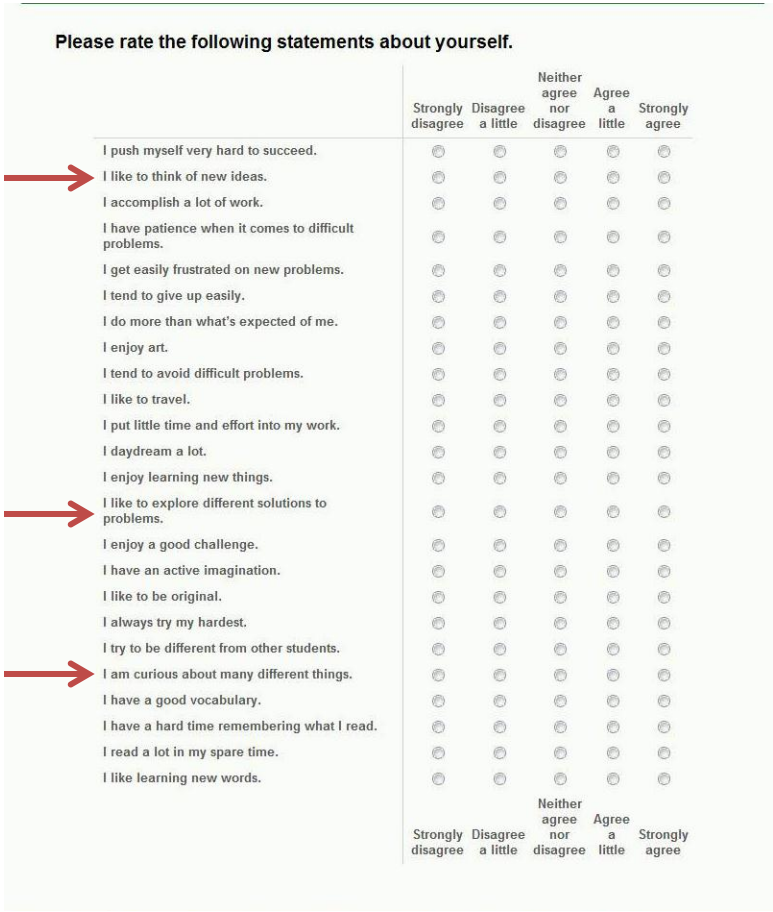
- Compare three stealth assessment estimates to validated external measures.
- Correlations should be reasonably high—but not too high as the "standard" measures are fairly limited.

Validation


Creativity
+
Openness self-report
+
Odd word test (home-made)
+
Wallach and Kogan's alternate uses test (1965)
ns
Conscientiousness stealth assessment

External Measures

Openness self-report



Please rate the following statements about yourself.

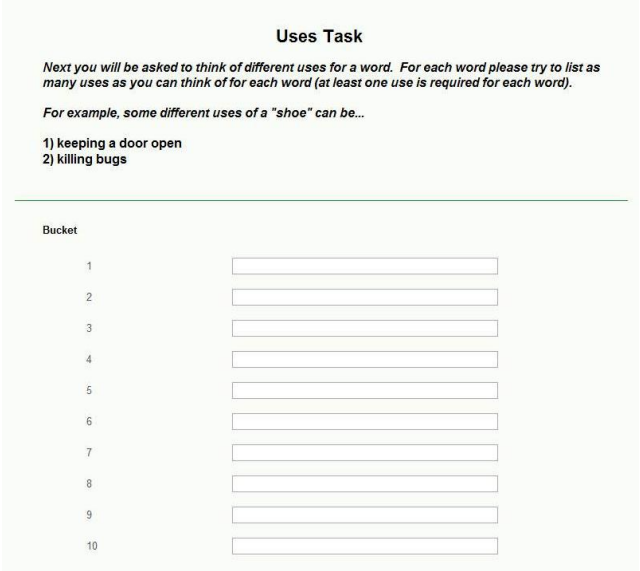
	Strongly disagree	Disagree a little	Neither agree nor disagree	Agree a little	Strongly agree
I push myself very hard to succeed.	○	○	○	○	○
I like to think of new ideas.	○	○	○	○	○
I accomplish a lot of work.	○	○	○	○	○
I have patience when it comes to difficult problems.	○	○	○	○	○
I get easily frustrated on new problems.	○	○	○	○	○
I tend to give up easily.	○	○	○	○	○
I do more than what's expected of me.	○	○	○	○	○
I enjoy art.	○	○	○	○	○
I tend to avoid difficult problems.	○	○	○	○	○
I like to travel.	○	○	○	○	○
I put little time and effort into my work.	○	○	○	○	○
I daydream a lot.	○	○	○	○	○
I enjoy learning new things.	○	○	○	○	○
I like to explore different solutions to problems.	○	○	○	○	○
I enjoy a good challenge.	○	○	○	○	○
I have an active imagination.	○	○	○	○	○
I like to be original.	○	○	○	○	○
I always try my hardest.	○	○	○	○	○
I try to be different from other students.	○	○	○	○	○
I am curious about many different things.	○	○	○	○	○
I have a good vocabulary.	○	○	○	○	○
I have a hard time remembering what I read.	○	○	○	○	○
I read a lot in my spare time.	○	○	○	○	○
I like learning new words.	○	○	○	○	○
	Strongly disagree	Disagree a little	Neither agree nor disagree	Agree a little	Strongly agree

Odd word test

grass flower moss cactus

	reason 1	reason 2	reason 3	reason 4	reason 5
grass, only one that					
flower, only one that					
moss, only one that					
cactus, only one that					

Wallach and Kogan's alternate uses test



Uses Task

Next you will be asked to think of different uses for a word. For each word please try to list as many uses as you can think of for each word (at least one use is required for each word).

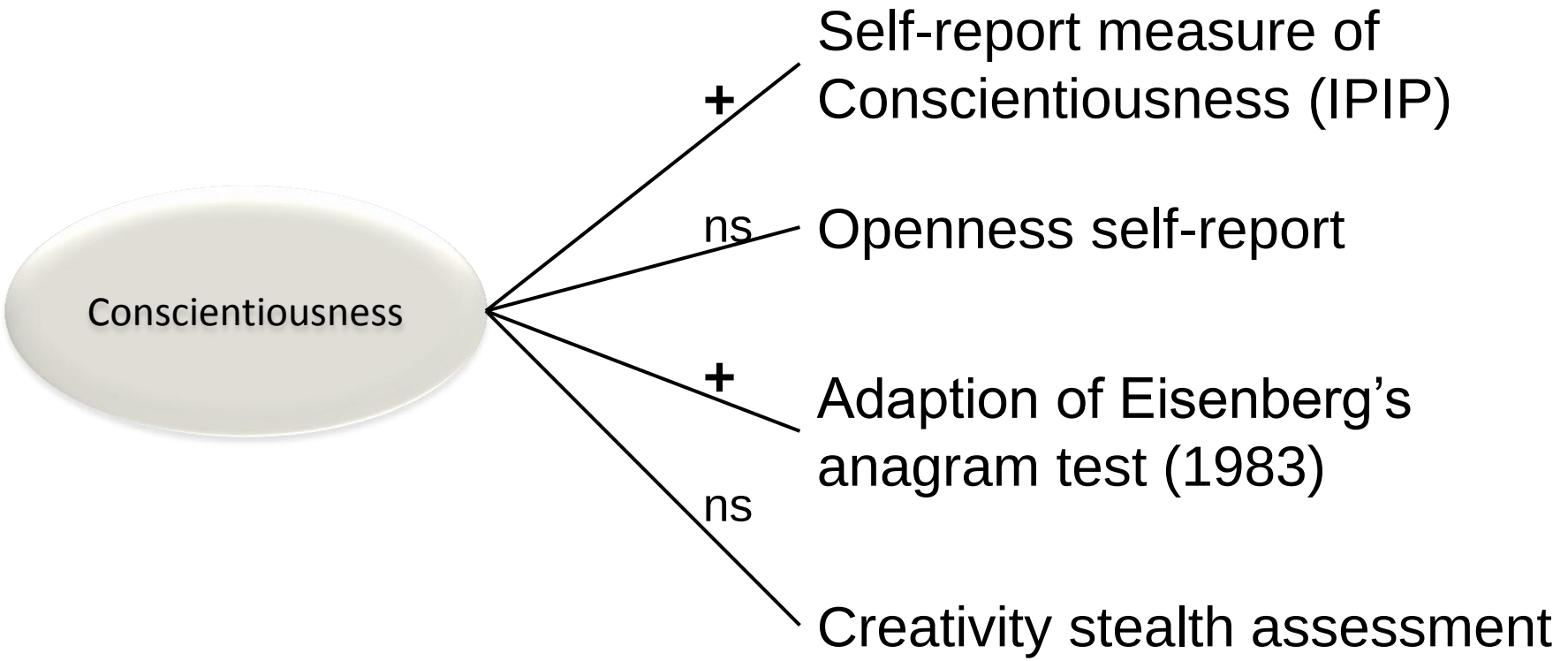
For example, some different uses of a "shoe" can be...

1) keeping a door open
2) killing bugs

Bucket

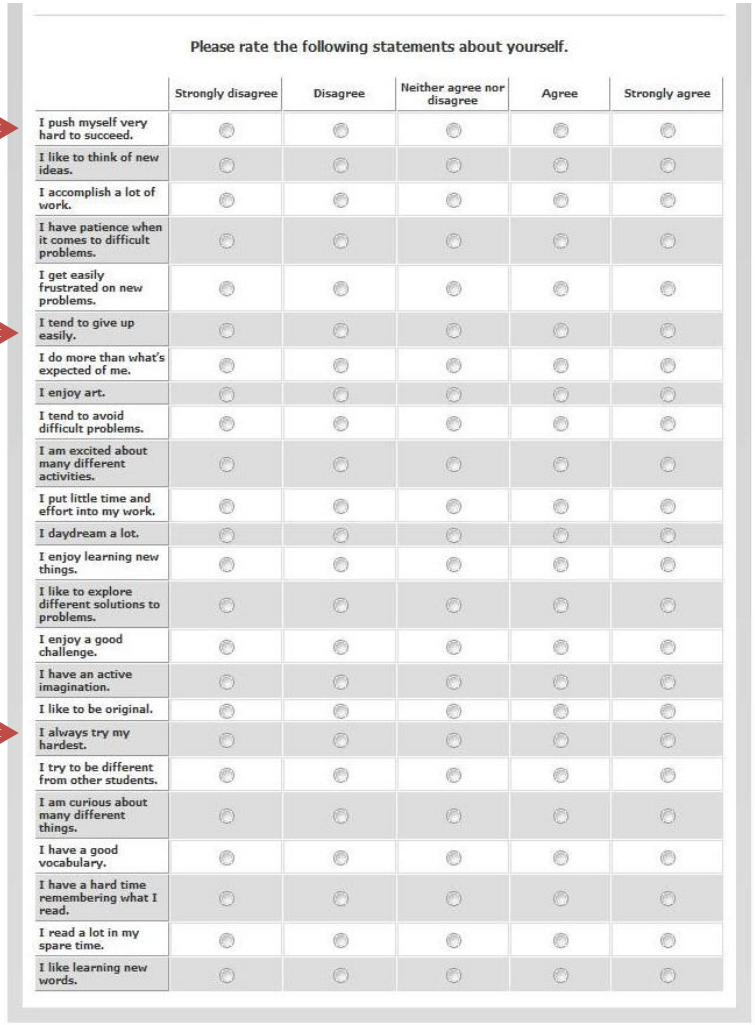
1
2
3
4
5
6
7
8
9
10

Validation


Conscientiousness
+
Self-report measure of Conscientiousness (IPIP)
ns
Openness self-report
+
Adaption of Eisenberg's anagram test (1983)
ns
Creativity stealth assessment

External Measures

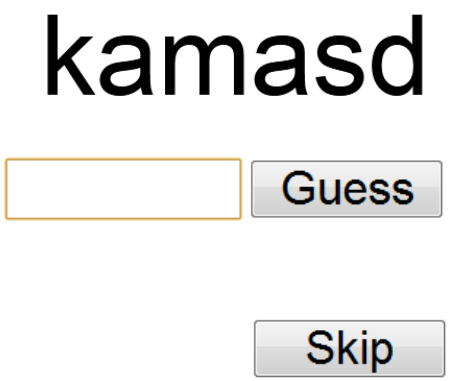
Self-report measure of conscientiousness



Please rate the following statements about yourself.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I push myself very hard to succeed.					
I like to think of new ideas.					
I accomplish a lot of work.					
I have patience when it comes to difficult problems.					
I get easily frustrated on new problems.					
I tend to give up easily.					
I do more than what's expected of me.					
I enjoy art.					
I tend to avoid difficult problems.					
I am excited about many different activities.					
I put little time and effort into my work.					
I daydream a lot.					
I enjoy learning new things.					
I like to explore different solutions to problems.					
I enjoy a good challenge.					
I have an active imagination.					
I like to be original.					
I always try my hardest.					
I try to be different from other students.					
I am curious about many different things.					
I have a good vocabulary.					
I have a hard time remembering what I read.					
I read a lot in my spare time.					
I like learning new words.					

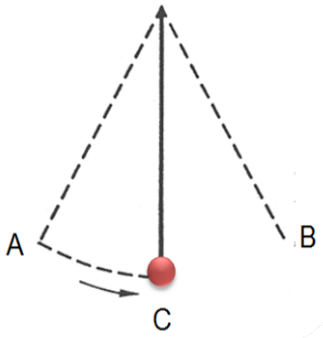
Adaption of Eisenberg's anagram test (1983)



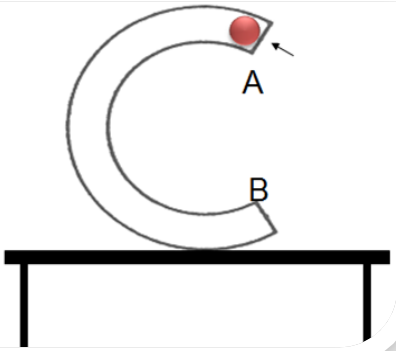
Learning

- In pretest-posttest design, we'll use the Force Concept Inventory (FCI), Masson et al. (2010) items, etc. to ***test conceptual understanding of physics concepts*** .
- MC items and constructed responses—e.g., applying impulses to objects, interpreting kinematics (i.e., position, velocity, and acceleration) in dot trace drawings.

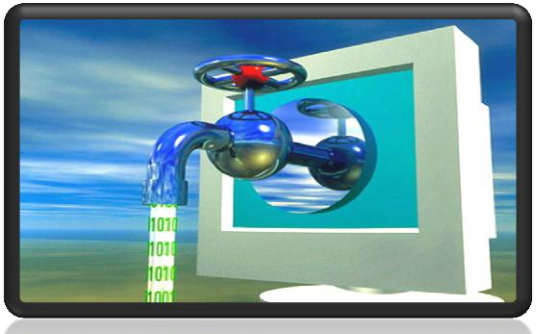
A metal ball is suspended by a string. The ball and the string move in an arc from point A to point B. While the ball is in motion, the string is cut at point C. Please draw the path that the ball would take *after* the string is cut.



Here is a hollow, circular tube mounted upright on a table top. You're looking at a side view of the table. A ball is shot into the end of the tube at point A. The ball leaves the other end of the tube at point B at high speed. Draw the path that the ball takes after it exits the tube.



Stealth Assessment – Nutshell



Extracts ongoing data from a user during gameplay, which provides the basis for making valid (evidence-based) inferences of competency states—at any time and at any grain size.

Captured data is converted to information that may be immediately and directly acted on (e.g., feedback, diagnostics and targeted support, changes to the game/environment).

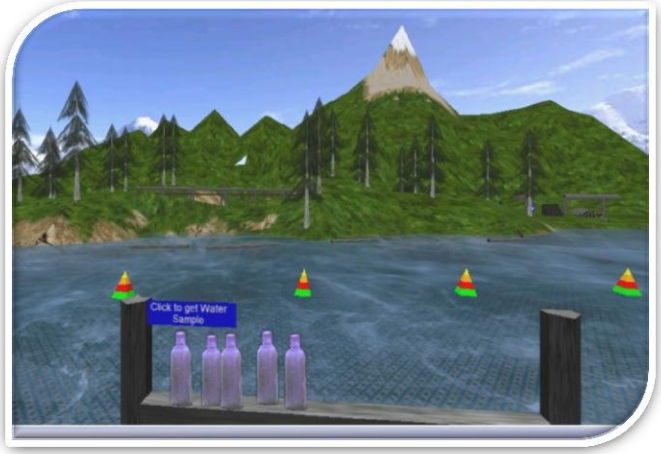
Components of stealth assessment may be reused in other systems (e.g., learning, training, or other gaming environments) thus substantially reducing costs.

Wrapping it Up

- Preparing students to succeed in 21st century requires supporting new competencies (e.g., *ability to solve complex problems*).

- There's also a need to design and develop assessments that are valid and reliable to meet these new educational needs (e.g., *ECD and stealth assessment*).

- Immersive games are ***fun/engaging***, and enable learning within complex, realistic, and relevant envir's. Also lead to social negotiation, ownership of learning, etc.



Wrapping it Up

- Bayes nets can be used in various ways to *improve learning* and performance.
  - Continuously gather evidence for *accurate, real-time estimates of comp's.*
  - Info on competencies used by (a) ***teachers*** (to adjust instruction & give good feedback), (b) ***system*** (to select new gaming experiences), and/or (c) ***students*** (to reflect on how they're doing).
- Current estimates of competency levels can be integrated into the game and displayed as *progress indicators.*
- ***This elevates valued competencies to the same level as health and weapons!***

Summary

- To address educational challenges and harness potential of immersive games, I presented an ECD-inspired idea re:
  - **Specifying competencies to be acquired from the game**
  - **Defining EMs that link game behaviors to competencies**
  - **Updating the learner model regularly**

- Using ECD, stealth assessment, and automated data collection and analysis tools is meant to collect valid evidence of students' emerging competencies, and *reduce teachers' workload* allowing them to focus on fostering student learning.

- Next steps: adapt content (feedback, difficulty levels, scaffolding, etc.) in game to fit current needs of player, test plug 'n play of models, evaluate learning effects (*n* = 200 middle school students), etc.

Poll

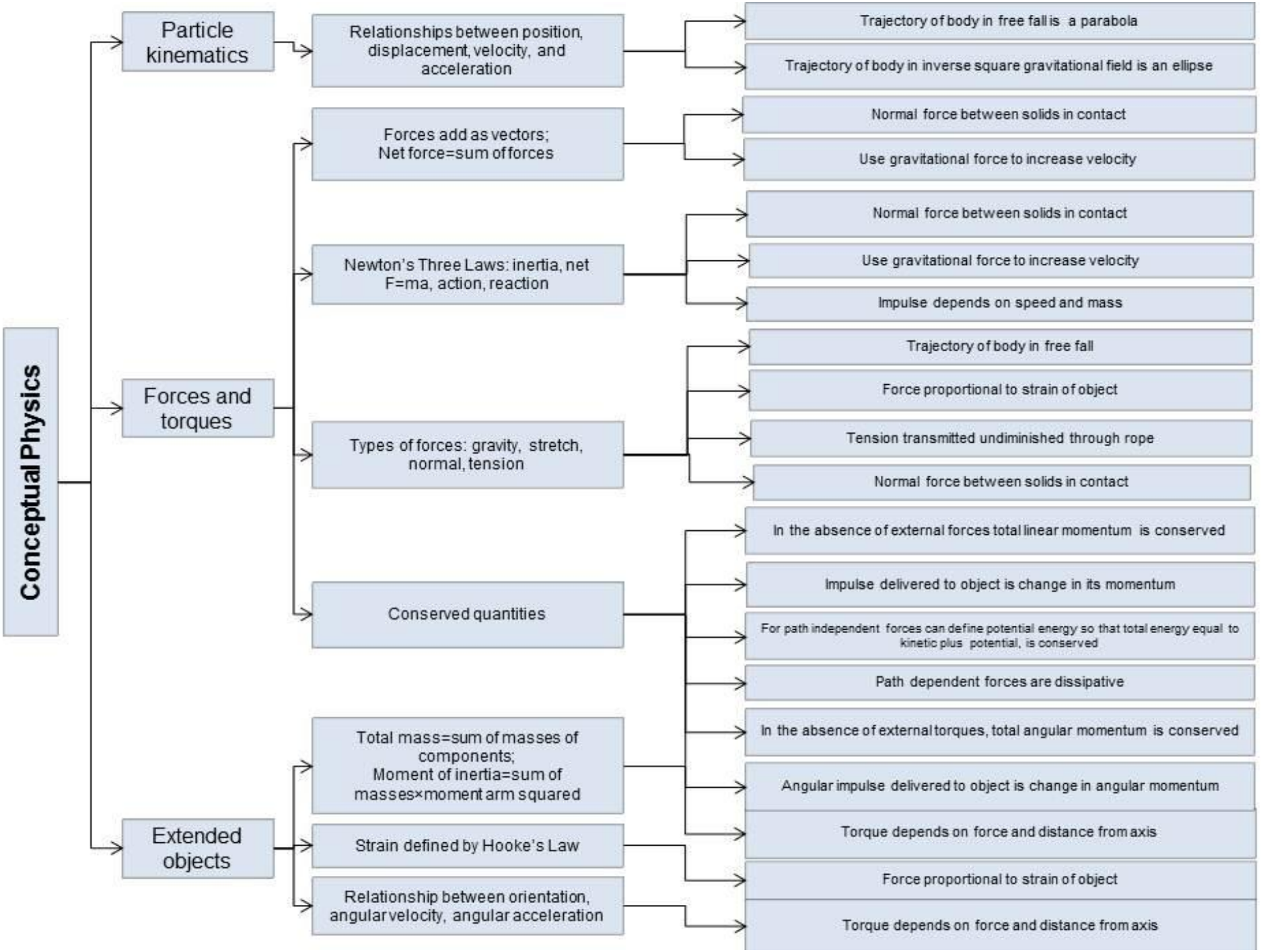
I’m thinking about changing the name from stealth to intrinsic assessment.

- Prefer the term STEALTH ASSESSMENT?
- Prefer the term INTRINSIC ASSESSMENT?

Thank you!

Questions?

Email: *vshute@fsu.edu*
Website: *http://www.myweb.fsu.edu/vshute*


Conceptual Physics
Particle kinematics
Relationships between position, displacement, velocity, and acceleration
Trajectory of body in free fall is a parabola
Trajectory of body in inverse square gravitational field is an ellipse
Forces and torques
Forces add as vectors; Net force=sum of forces
Normal force between solids in contact
Use gravitational force to increase velocity
Newton's Three Laws: inertia, net F=ma, action, reaction
Normal force between solids in contact
Use gravitational force to increase velocity
Impulse depends on speed and mass
Types of forces: gravity, stretch, normal, tension
Trajectory of body in free fall
Force proportional to strain of object
Tension transmitted undiminished through rope
Normal force between solids in contact
Conserved quantities
In the absence of external forces total linear momentum is conserved
Impulse delivered to object is change in its momentum
For path independent forces can define potential energy so that total energy equal to kinetic plus potential, is conserved
Path dependent forces are dissipative
Extended objects
Total mass=sum of masses of components; Moment of inertia=sum of masses×moment arm squared
In the absence of external torques, total angular momentum is conserved
Angular impulse delivered to object is change in angular momentum
Strain defined by Hooke's Law
Torque depends on force and distance from axis
Force proportional to strain of object
Relationship between orientation, angular velocity, angular acceleration
Torque depends on force and distance from axis