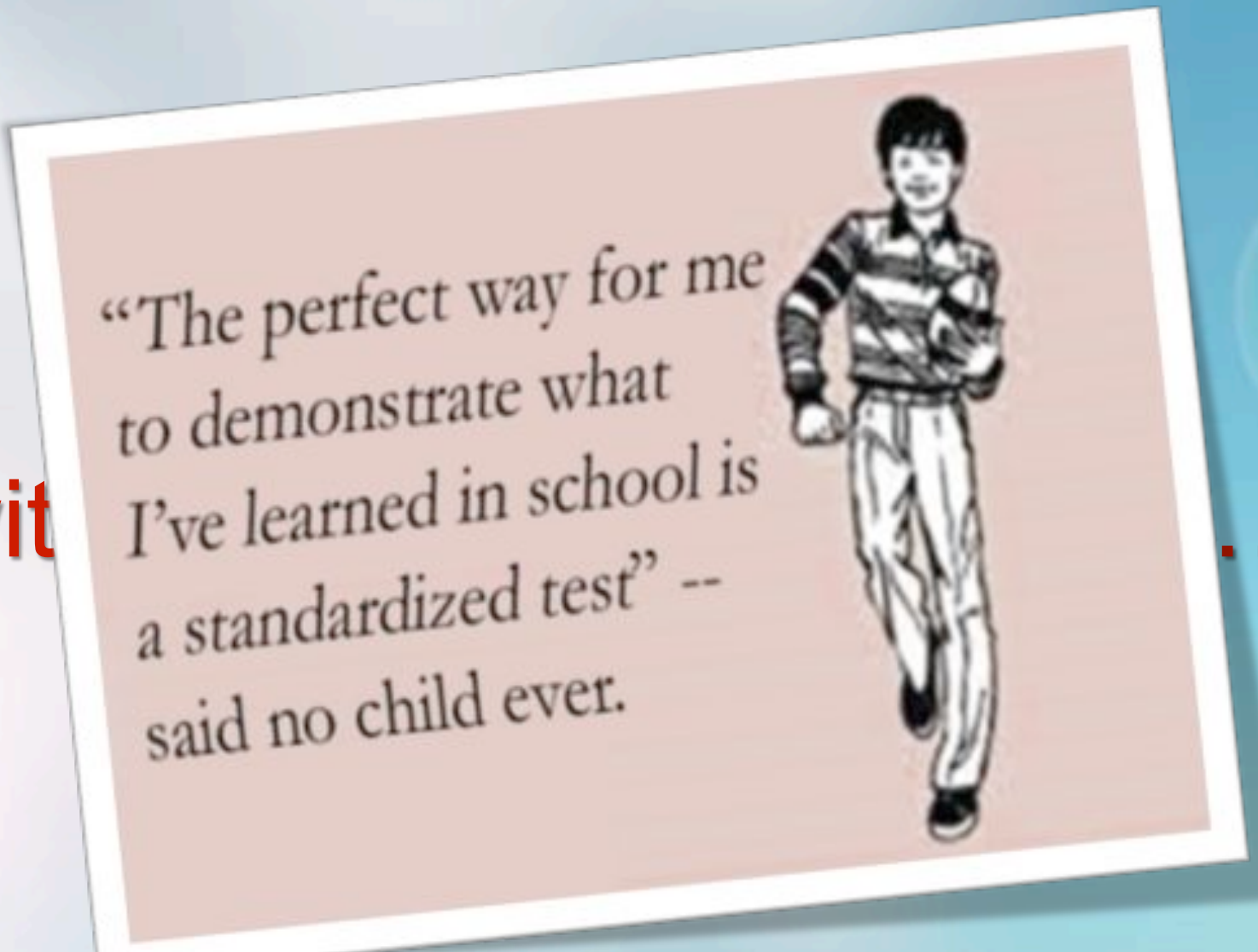
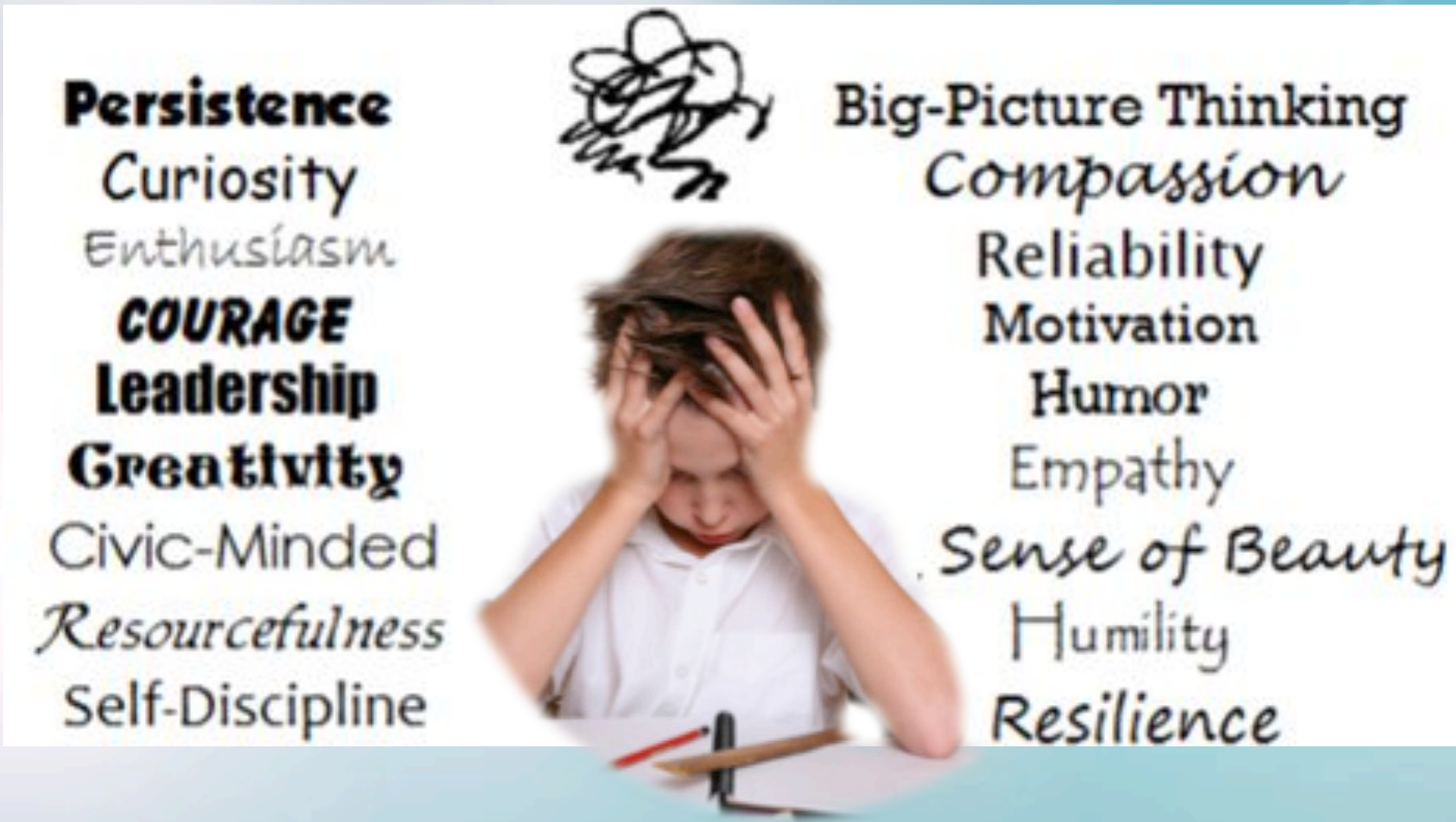
wit



. .

Val Shute, FSU (NSF/CRA Big Data Workshop, 6/2/15)

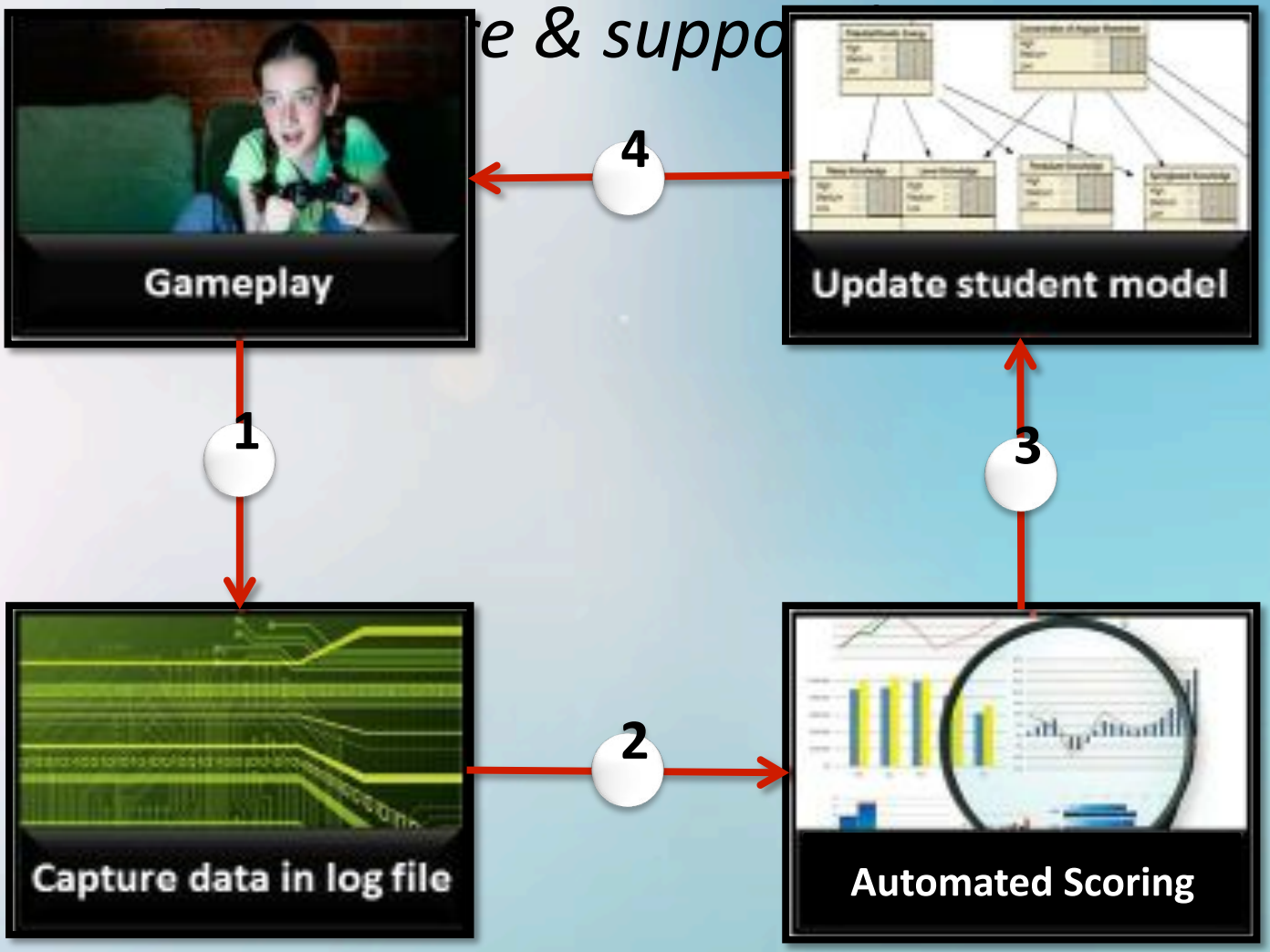
What Standardized Tests Don't Measure



SEE: Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M-W. (in press). Advances in the science of assessment. To appear in *Educational Assessment*.

Stealth Assessment in Games:

e & suppo



Stealth Assessment Features

Seamless & Ubiquitous

When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative.

Formative, not Summative

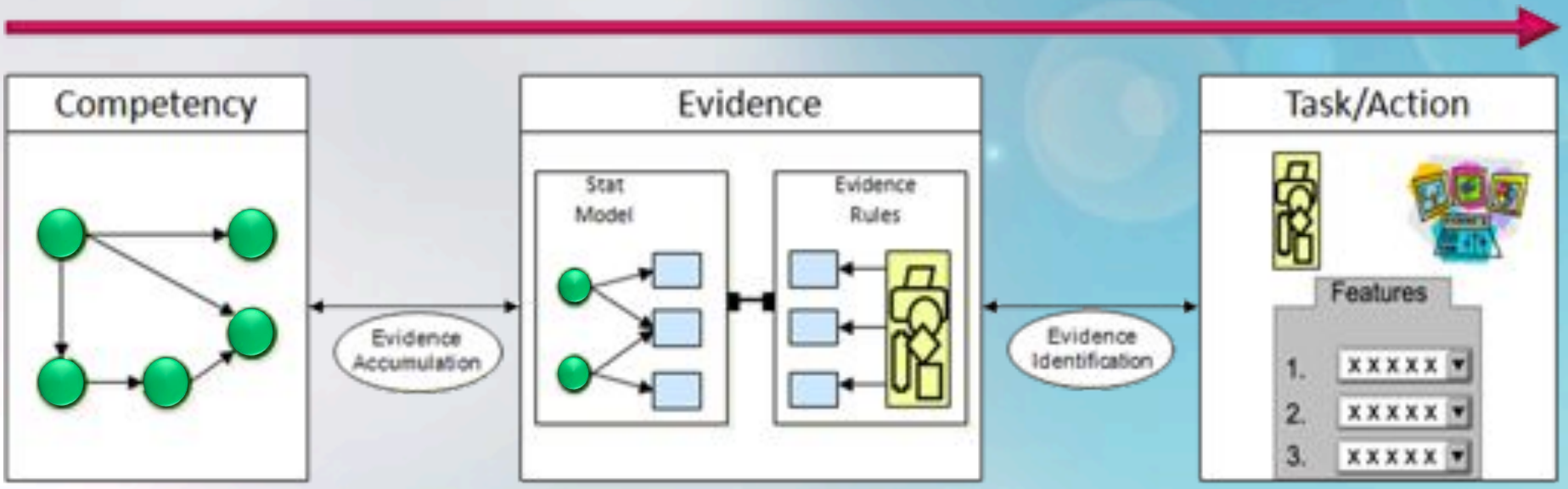
Accurate & Rich Learner Models

Invisible assessment, transparent support!

ECD

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

Assessment Models & Metrics

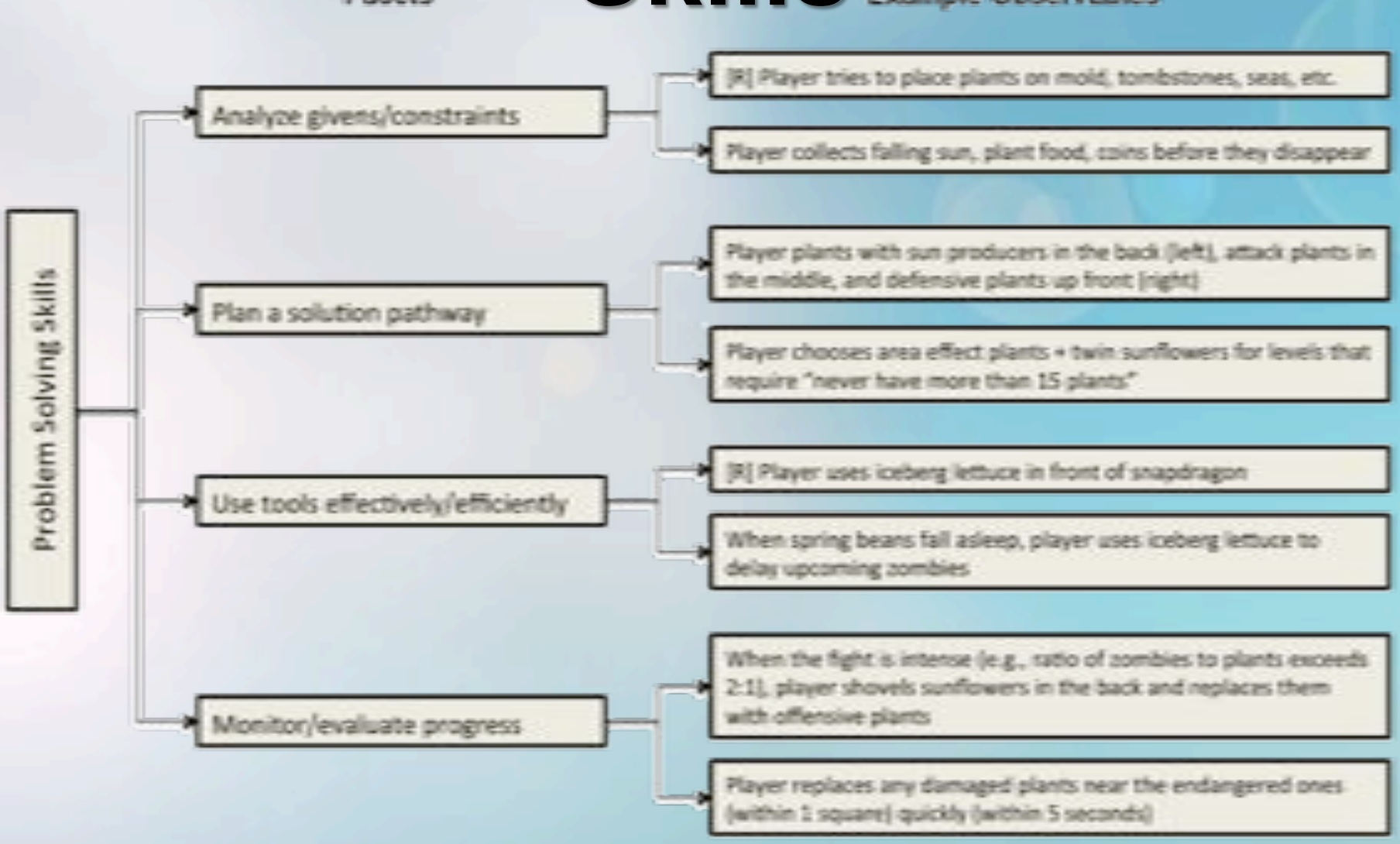


Monitor & Diagnose Success

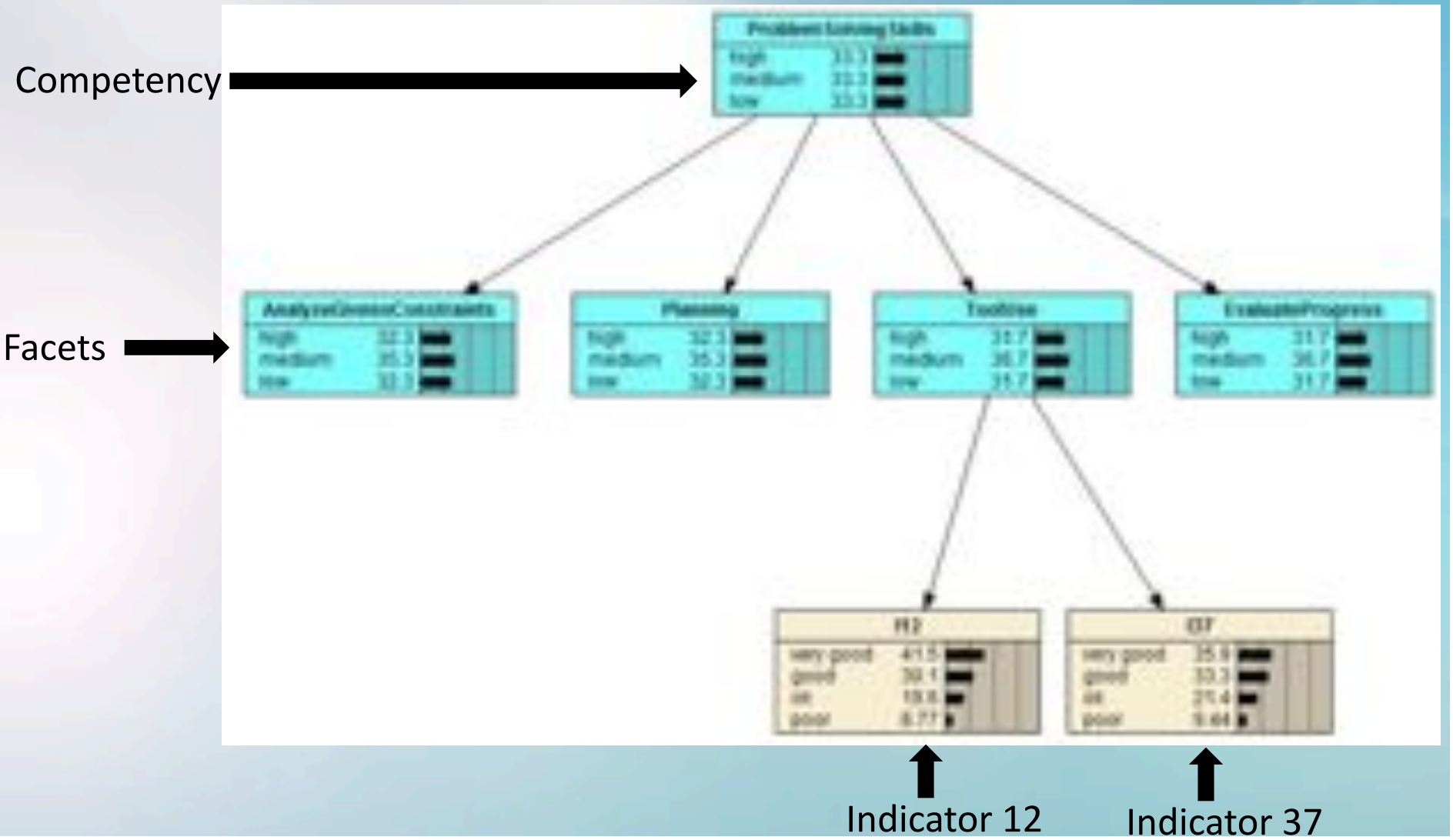
Current Work

Plants vs. Zombies 2 (measuring & supporting problem solving skills)

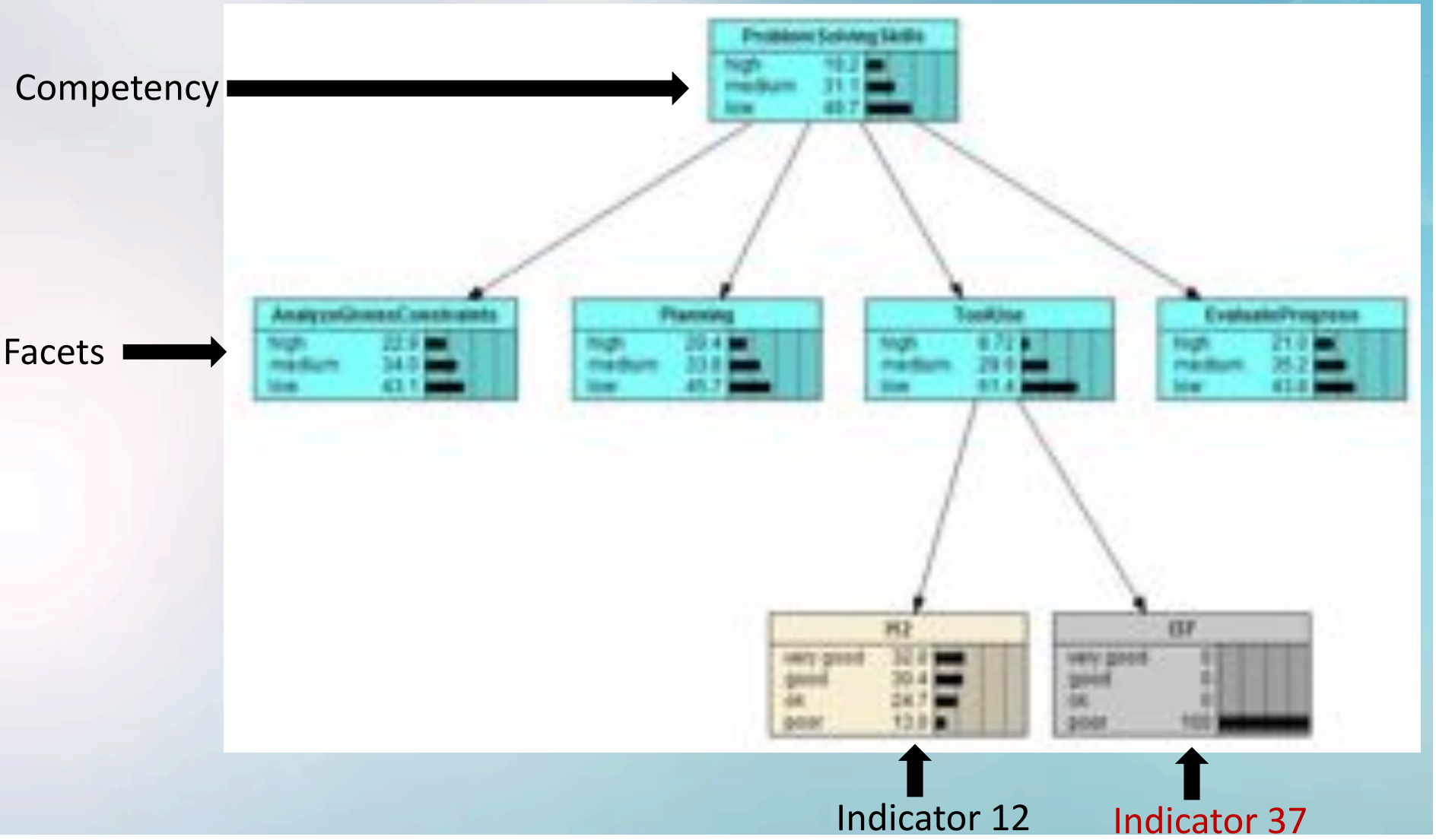
CM Problem Solving Skills



Problem Solving BN (priors)



Problem Solving BN (data)



Preliminary Validity Results

Despite the small sample size, age, and SES of participants (47 students, 7th grade, low SES):

1) Problem solving in PvZ significantly correlates with **Raven's** ($r = .40$, $p < .01$).
2) Problem solving in PvZ significantly correlates with **MicroDYN** ($r = .48$, $p < .01$)
3) These preliminary findings suggest that our stealth assessment estimates are valid, and we are currently using machine learning in *netica* to improve the BNs.

Thank you!

Questions?

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

Telemetry Hooks

Hook Name	Additional Data	Description	Rationale
Action_shovel 	type : string xCoordinate : int yCoordinate : int	Recorded when the player removes a plant from the battlefield. Included in the telemetry -- type of plant (e.g. Sunflower, Pea Shooter, Wall Nut, etc.) and the coordinates from where it was removed.	This is a core mechanic of the game providing insight into player's strategy and decision making –*use tools appropriately* and *monitor/evaluate progress*.

Augmented Q-Matrix

Indicator		Facets					
#	Description	Analyze givens, constraints (F1)	Analyze goals (F2)	Plan a solution pathway (F3)	Effective, efficient tools/resources use (F4)	Monitor, evaluate progress (F5)	levels of c relationships
	1) [E] Players attempted to place plants on top of mulch	1	0	0	0	0	4 NA
	3 Plant sunflowers close to home base (last two columns)	1	0	0	1	0	4 NA
9, 10	Players choose cabbage pull, green beans, and/or	1	0	0	1	0	4 conjunctive
9, 11	Players choose area of effort plants/ tools sunflower for	0	1	1	0	0	4 conjunctive
8,20,23,40	Players choose correct carrots and/or cherry tomat for	0	1	1	0	0	2 conjunctive
	7) Collect falling sun before it disappears (ratio #-	1	0	0	0	0	4 NA
	8) Plant 1-3 sunflowers at the beginning before the second	1	0	0	0	0	2 conjunctive
	9) collect plant food before it disappears (ratio #	1	0	0	0	0	4 NA
	12) Use plants before the conveyor belt is full	1	0	0	0	0	4 NA
13,36	[E] Use plant food when there is < 3 zombies on the	1	0	0	1	0	4 conjunctive
13,25	Use plant food to take down > 5 zombies in the yard or	1	0	0	1	0	4 conjunctive
14,28	[E] Use powers when there is > 3 zombies on the screen	1	0	0	1	0	4 conjunctive
13,27	Use powers to take down >5 zombies in the yard or still	1	0	0	1	0	4 conjunctive
	16) Players plant high damage plants in rows with many	0	1	0	0	0	4 NA
	17 Players plant tough plants in rows with >3 zombies, fast	0	1	0	0	0	4 NA
	18) Players plant with sun producers (i.e., sunflower, twin	0	0	1	0	0	4 NA
34, 42	Players choose twin sunflowers/sun plant food to	0	0	1	0	0	2 conjunctive
	29) Use the shovel to remove a weaker plant (e.g.,	1	0	0	1	0	2 conjunctive
	30) Plant iceberg lettuce/walnut within 2 squares in front to	0	0	0	1	0	2 NA
	31) Use coconut cannon to kill >3 zombies at a time	0	0	0	1	0	4 NA
	32 Plant bonk choy one square to the right of a zombie to kil	0	0	0	1	0	2 NA
	33) Calculate ratio of successfully "armed" potato mines wil	0	0	0	1	0	4 NA
	35) Players choose iceberg lettuce for their plant inventory if	0	0	1	1	0	2 conjunctive
	36) Players choose spikeweed or spikerock for their plant	0	0	1	1	0	2 conjunctive
	37) [E] Players plant iceberg lettuce within range of a	0	0	0	1	0	4 NA
	38) Players plant spring bean in a square immediately	0	0	0	1	0	4 NA
	41) In levels where the sun resource is limited (e.g. only	0	0	0	1	0	4 NA
	43) # of replaced plants within 3 seconds at the same	0	0	0	0	1	4 NA
	44 When the right a cannon (e.g. ratio of zombies to plants	0	0	0	0	1	2 NA
	45) Improvement on plant layout (#%) after a fail	0	0	0	0	1	4 NA
	46) Improvement on sequence of plants planted (#%) after fail	0	0	0	0	1	4 NA

Dashboard Rules

- Node is grey ("need more evidence") if:
  - calculate 3 absolute values: |p(high)-p(low)|, |p(high)-p(med)|, |p(med)-p(low)|
  - any 2(of 3) values is ≤ .15, then node for the competency is grey
- If the node is *not* grey, calculate EAP values, which is represented by p(high)-p(low) (range from -1 to 1)
  - Node color:

    Green, if EAP falls in [0.34, 1]

    Yellow, if EAP falls in [-0.34, 0.33]

    Red, if EAP falls in [-1, -0.33]

Taxonomy of Plants

Plant	Image	Description	Unlocked	Plant Food Ability	Sun Cost	Recharge
Sunflower		Sunflowers are essential for you to produce extra sun. Try planting as many as you can!	After beating Tutorial Level - Day 1	Produces 150 sun	50 sun	Fast
Peashooter		Peashooters are your first line of defense. They shoot peas at attacking zombies.	Available at the beginning of the game	Turns them into faster shooting Gatling Peashooting. Mf peas for three seconds.	100 sun	Fast
Wall-nut		Wall-nuts have hard shells which you can use to protect other plants.	After beating Player's House - Day 2	Gives it a hard armor shell for extra strength.	50 sun	Sluggish
Potato Mine		Potato Mines explode on contact, but they take time to arm. Plant them ahead of zombies.	After beating Player's House - Day 3	Arms itself and produces two armed Potato Mines on other spaces.	25 sun	Sluggish
Boomerang		Boomerangs can hit up to three targets in their lane, twice!	After beating Ancient Egypt - Day 1	Shoots approximately ten boomerangs up, down, left and right.	175 sun	Fast
Cabbage-pult		Cabbage-pults hurl cabbages at the enemy.	After beating Ancient Egypt - Day 2	Launches several cabbages at once, hitting every zombie onscreen.	100 sun	Fast

...

Taxonomy of Zombies

Name	Image	Toughness	Health	Speed	Special
[illegible]		Average	[illegible]	Basic	
[illegible]		Protected	[illegible]	Basic	
[illegible]		Hardened	[illegible]	Basic	
[illegible]		Average	[illegible]	Basic	[illegible]
[illegible]		Average	[illegible]	Hungry	[illegible]
[illegible]		Great	[illegible]	Hungry	[illegible]

...

Developmental Review 32 (2012) 125–160

Contents lists available at SciVerse ScienceDirect

Developmental Review

journal homepage: www.elsevier.com/locate/dr

Review

Children's scientific curiosity: In search of an operational definition of an elusive concept ☆

Jamie Jirout [a,*], David Klahr [b]

[a] Department of Psychology, Temple University, United States
[b] Department of Psychology, Carnegie Mellon University, United States

ARTICLE INFO

Article history:
Received 20 June 2011
Revised 31 March 2012
Available online 26 April 2012

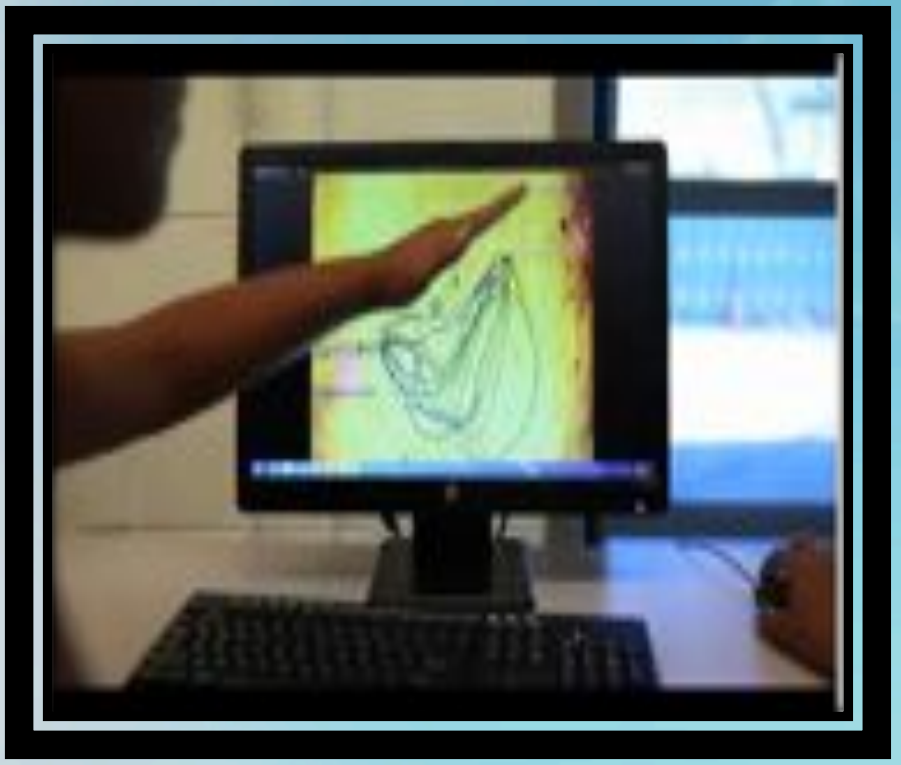
Keywords:
Curiosity
Children
Scientific thinking
Exploratory behavior

ABSTRACT

Although curiosity is an undeniably important aspect of children's cognitive development, a universally accepted operational definition of children's curiosity does not exist. Almost all of the research on measuring curiosity has focused on adults, and has used predominately questionnaire-type measures that are not appropriate for young children. In this review we (a) synthesize the range of definitions and measures of children's curiosity and (b) propose a new operational definition and measurement procedure for assessing and advancing scientific curiosity in young children. In the first part of the paper, we summarize Loewenstein's (1994) review of theoretical perspectives on adult curiosity, and critically evaluate a wide range of efforts to create and implement operational measures of curiosity in children. In the second part, we return to Loewenstein's theory and present an argument for adopting his "information-gap" theory of curiosity as a framework for reviewing various procedures that have been suggested for measuring children's exploratory curiosity. Finally, we describe a new paradigm for measuring exploratory curiosity in preschool children, defining curiosity as the threshold of desired uncertainty in the environment that leads to exploratory behavior. We present data demonstrating the reliability and validity of this measure, discuss initial results on developmental differences in young children's curiosity, and conclude with a general summary and suggestions for future research.

Stealth Assessment

- Allows us to:
  - Extract *ongoing* info from a learner (across disparate sources)
  - Make *accurate inferences* of competencies
  - React in *immediate* & helpful ways.



- Accomplished via automated scoring and machine-based reasoning techniques.

Issues with Standardized Tests

There's growing criticism of large-scale achievement tests (e.g., Sackett et al., 2008; Zwick, 2004). For example, these tests:

- predict badly
- do not measure all the relevant determinants of important criteria related to achievement and learning
- are subject to coaching
- do not measure genuine ability and classroom achievement
- are biased against members of racial and ethnic minority groups
- are subject to motivational differences among students
- function largely as "class" or "wealth" tests because learners from affluent or high socio-economic backgrounds tend to perform better on the tests than learners from disadvantaged or low socio-economic backgrounds.

See: Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M-W. (in press). Advances in the science of assessment. To appear in *Educational Assessment*.

Wrapping it Up

- Preparing kids to succeed in 21st century requires supporting new competencies—thus a need to develop assessments that are valid & (e.g., *ECD and stealth assessment*).

- Use immersive games are ***fun/ engaging***, and enable learning within complex, realistic, and relevant environments.