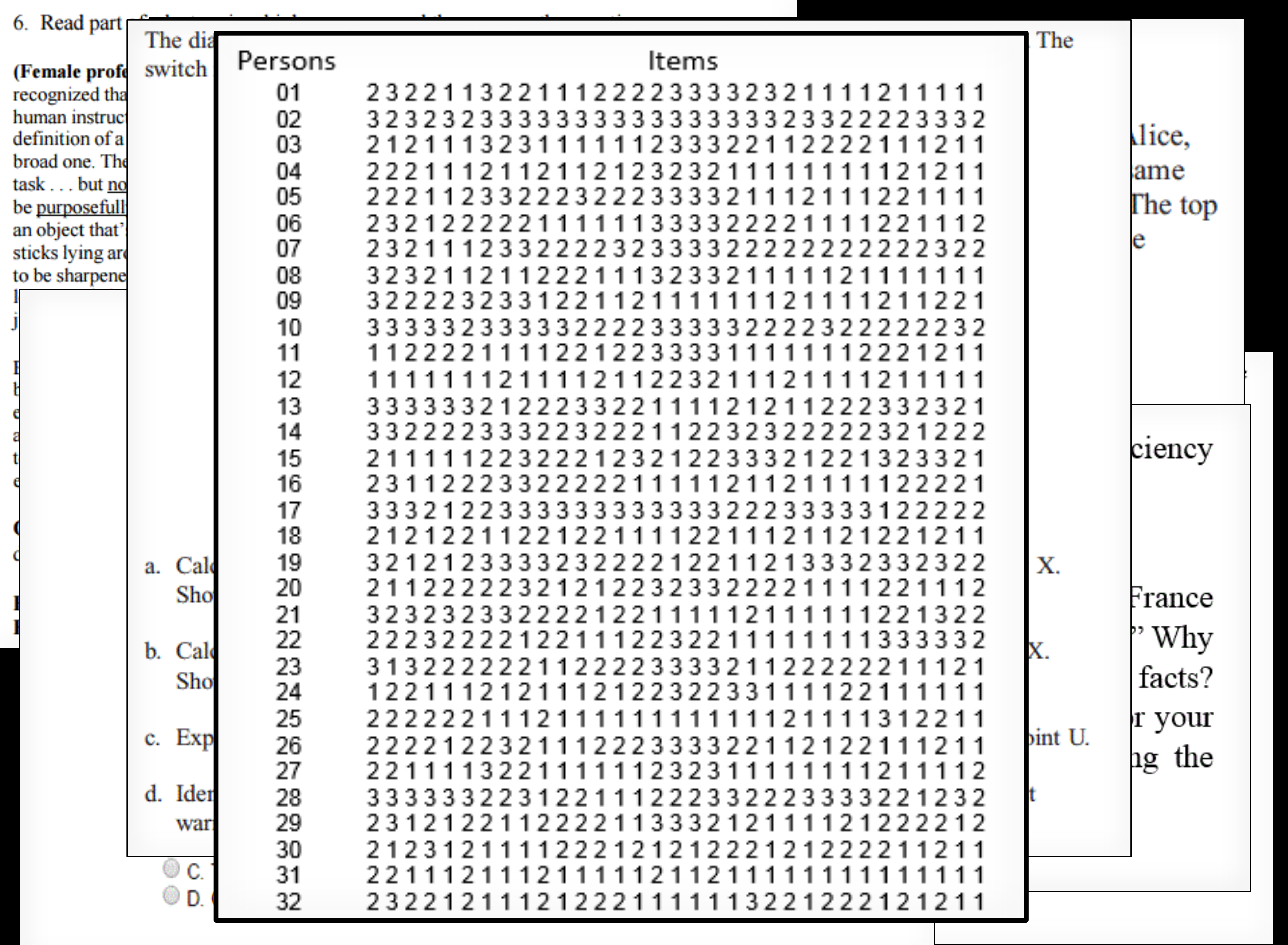
On Integrating Psychometrics and Learning Analytics in Complex Assessments

Robert J. Mislevy

Educational Testing Service

The Sixteenth Annual Maryland Conference:
Data Analytics and Psychometrics: Informing Assessment Practices.

Maryland Assessment Research Center, University of Maryland,
College Park, MD, November 3-4, 2016



6. Read part

(Female profe
recognized tha
human instruct
definition of a
broad one. The
task . . . but no
be purposefull
an object that'
sticks lying are
to be sharpene

The dia
switch

a. Calc
Sho

b. Calc
Sho

c. Exp

d. Iden
war

○ C.
○ D.

The

lice,
ame
The top
e

ciency

X.

X.

oint U.

t

France
" Why
facts?
r your
ng the

Persons	Items																																
01	2	3	2	2	1	1	3	2	2	1	1	1	2	2	2	2	3	3	3	3	2	3	2	1	1	1	1	2	1	1	1	1	1
02	3	2	3	2	3	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	3	2	2	2	2	3	3	3	2
03	2	1	2	1	1	1	3	2	3	1	1	1	1	1	1	2	3	3	3	2	2	1	1	2	2	2	2	1	1	1	2	1	1
04	2	2	2	1	1	1	2	1	1	2	1	1	2	1	2	3	2	3	2	1	1	1	1	1	1	1	1	1	2	1	2	1	1
05	2	2	2	1	1	2	3	3	2	2	2	3	2	2	2	3	3	3	3	2	1	1	1	2	1	1	1	2	2	1	1	1	1
06	2	3	2	1	2	2	2	2	2	1	1	1	1	1	1	3	3	3	3	2	2	2	2	1	1	1	1	2	2	1	1	1	2
07	2	3	2	1	1	1	2	3	3	2	2	2	2	3	2	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	3	2	2
08	3	2	3	2	1	1	2	1	1	2	2	2	1	1	1	3	2	3	3	2	1	1	1	1	1	2	1	1	1	1	1	1	1
09	3	2	2	2	2	3	2	3	3	1	2	2	1	1	2	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	2	2	1
10	3	3	3	3	3	2	3	3	3	3	3	2	2	2	2	3	3	3	3	3	2	2	2	2	3	2	2	2	2	2	2	3	2
11	1	1	2	2	2	2	1	1	1	1	2	2	1	2	2	3	3	3	3	1	1	1	1	1	1	1	2	2	2	1	2	1	1
12	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1	2	2	3	2	1	1	1	2	1	1	1	1	2	1	1	1	1	1
13	3	3	3	3	3	3	2	1	2	2	2	3	3	2	2	1	1	1	1	2	1	2	1	1	2	2	2	3	3	2	3	2	1
14	3	3	2	2	2	2	3	3	3	2	2	3	2	2	2	1	1	2	2	3	2	3	2	2	2	2	2	3	2	1	2	2	2
15	2	1	1	1	1	1	2	2	3	2	2	2	1	2	3	2	1	2	2	3	3	3	2	1	2	2	1	3	2	3	3	2	1
16	2	3	1	1	2	2	2	3	3	2	2	2	2	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	2	2	2	2	1
17	3	3	3	2	1	2	2	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	3	3	3	3	1	2	2	2	2	2
18	2	1	2	1	2	2	1	1	2	2	1	2	2	1	1	1	1	2	2	1	1	1	2	1	1	2	1	2	2	1	2	1	1
19	3	2	1	2	1	2	3	3	3	3	2	3	2	2	2	2	1	2	2	1	1	2	1	3	3	3	2	3	3	2	3	2	2
20	2	1	1	2	2	2	2	2	3	2	1	2	1	2	2	3	2	3	3	2	2	2	2	1	1	1	1	2	2	1	1	1	2
21	3	2	3	2	3	2	3	3	2	2	2	2	1	2	2	1	1	1	1	1	2	1	1	1	1	1	1	2	2	1	3	2	2
22	2	2	2	3	2	2	2	2	1	2	2	1	1	1	2	2	3	2	2	1	1	1	1	1	1	1	1	3	3	3	3	3	2
23	3	1	3	2	2	2	2	2	2	1	1	2	2	2	2	3	3	3	3	2	1	1	2	2	2	2	2	2	1	1	1	2	1
24	1	2	2	1	1	1	2	1	2	1	1	1	2	1	2	2	3	2	2	3	3	1	1	1	1	2	2	1	1	1	1	1	1
25	2	2	2	2	2	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	3	1	2	2	1	1
26	2	2	2	2	1	2	2	3	2	1	1	1	2	2	2	3	3	3	3	2	2	1	1	2	1	2	2	1	1	1	2	1	1
27	2	2	1	1	1	1	3	2	2	1	1	1	1	1	1	2	3	2	3	1	1	1	1	1	1	1	1	2	1	1	1	1	2
28	3	3	3	3	3	3	2	2	3	1	2	2	1	1	1	2	2	2	3	3	2	2	2	3	3	3	3	2	2	1	2	3	2
29	2	3	1	2	1	2	2	1	1	2	2	2	2	1	1	3	3	3	2	1	2	1	1	1	1	2	1	2	2	2	2	1	2
30	2	1	2	3	1	2	1	1	1	1	2	2	2	1	2	1	2	1	2	2	2	1	2	1	2	2	2	2	1	1	2	1	1
31	2	2	1	1	1	2	1	1	1	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	2	3	2	2	1	2	1	1	1	2	1	2	2	2	1	1	1	1	1	1	3	2	2	1	2	2	2	1	2	1	2	1	1

The Standard Ed Measurement Paradigm

The aim is “measuring a construct,” framed in trait or behavioral psychology.

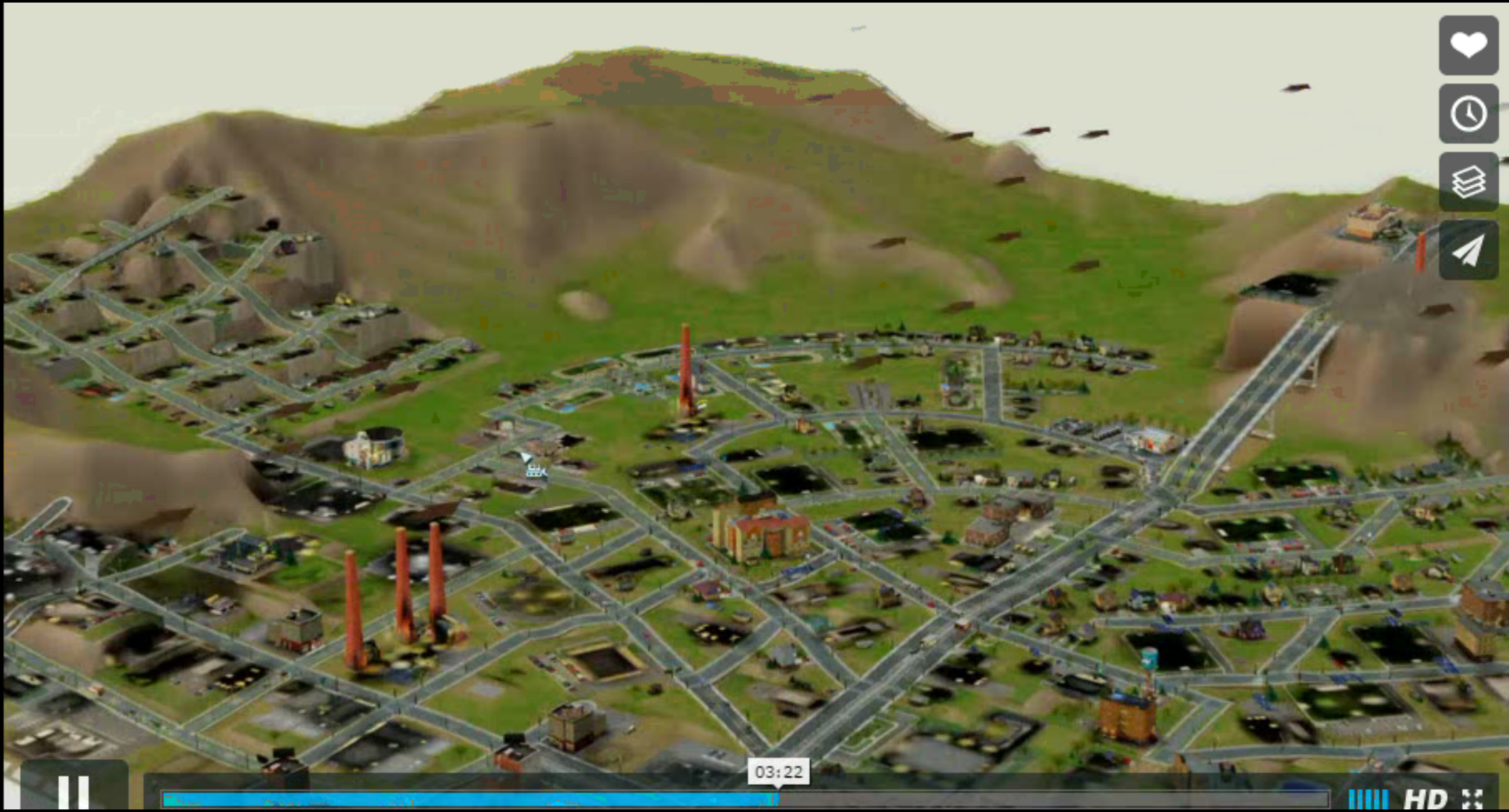
Usually a single measure is desired.

Each task (item) is a self-contained situation that evokes a response that provides evidence about the construct.

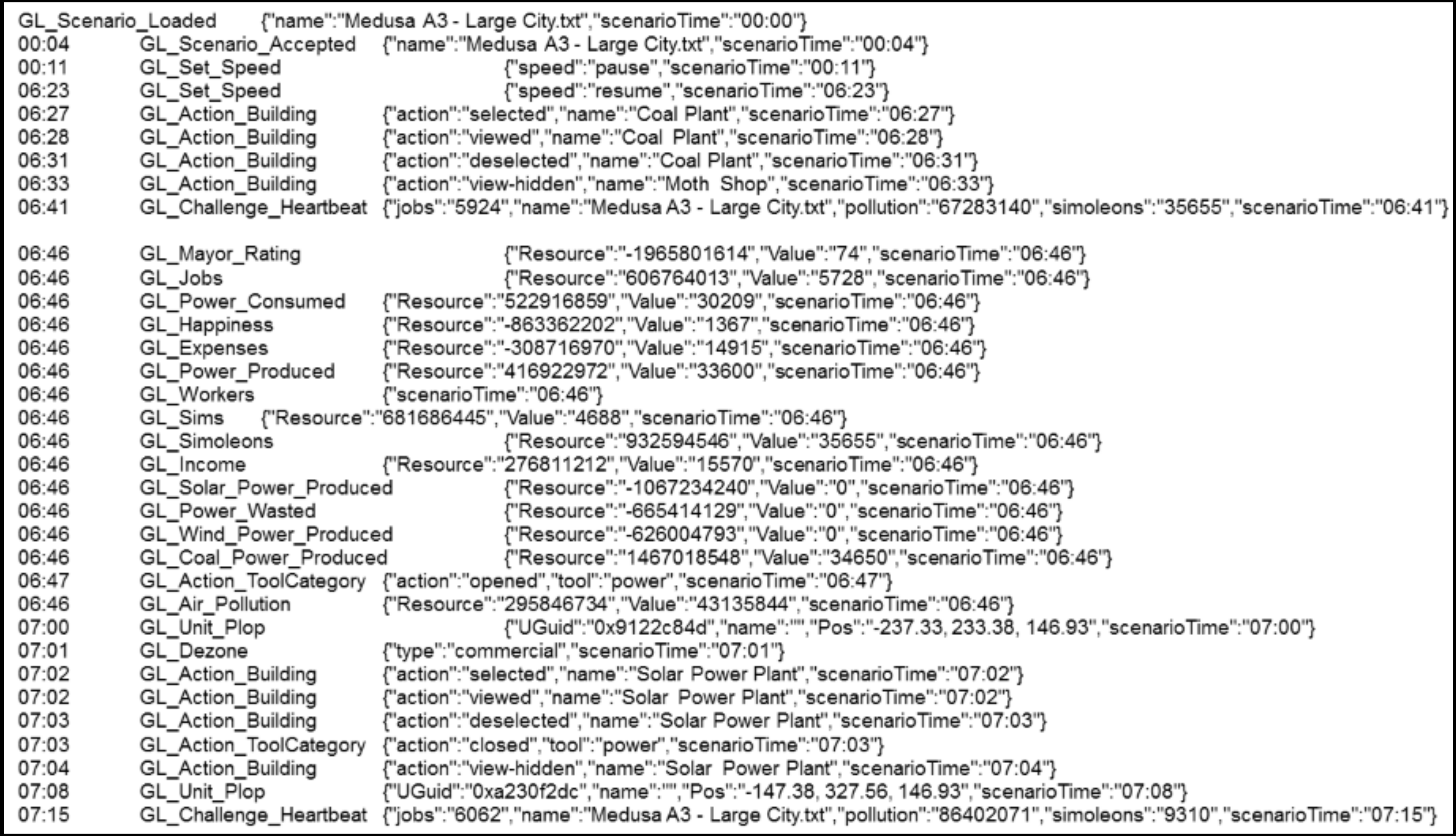
Each response is evaluated to provide an item score.

A test score accumulates evidence over items, often summing item scores, sometimes through a latent-variable model such as item response theory (IRT).

A snippet of SimCityEDU: Pollution Challenge!



A snippet of data from SimCityEDU



```
GL_Scenario_Loaded      {"name":"Medusa A3 - Large City.txt","scenarioTime":"00:00"}
00:04        GL_Scenario_Accepted    {"name":"Medusa A3 - Large City.txt","scenarioTime":"00:04"}
00:11        GL_Set_Speed                          {"speed":"pause","scenarioTime":"00:11"}
06:23        GL_Set_Speed                          {"speed":"resume","scenarioTime":"06:23"}
06:27        GL_Action_Building        {"action":"selected","name":"Coal Plant","scenarioTime":"06:27"}
06:28        GL_Action_Building        {"action":"viewed","name":"Coal  Plant","scenarioTime":"06:28"}
06:31        GL_Action_Building        {"action":"deselected","name":"Coal Plant","scenarioTime":"06:31"}
06:33        GL_Action_Building        {"action":"view-hidden","name":"Moth  Shop","scenarioTime":"06:33"}
06:41        GL_Challenge_Heartbeat  {"jobs":"5924","name":"Medusa A3 - Large City.txt","pollution":"67283140","simoleons":"35655","scenarioTime":"06:41"}

06:46        GL_Mayor_Rating                       {"Resource":"-1965801614","Value":"74","scenarioTime":"06:46"}
06:46        GL_Jobs                               {"Resource":"606764013","Value":"5728","scenarioTime":"06:46"}
06:46        GL_Power_Consumed       {"Resource":"522916859","Value":"30209","scenarioTime":"06:46"}
06:46        GL_Happiness              {"Resource":"-863362202","Value":"1367","scenarioTime":"06:46"}
06:46        GL_Expenses               {"Resource":"-308716970","Value":"14915","scenarioTime":"06:46"}
06:46        GL_Power_Produced       {"Resource":"416922972","Value":"33600","scenarioTime":"06:46"}
06:46        GL_Workers                {"scenarioTime":"06:46"}
06:46        GL_Sims     {"Resource":"681686445","Value":"4688","scenarioTime":"06:46"}
06:46        GL_Simoleons                          {"Resource":"932594546","Value":"35655","scenarioTime":"06:46"}
06:46        GL_Income                 {"Resource":"276811212","Value":"15570","scenarioTime":"06:46"}
06:46        GL_Solar_Power_Produced               {"Resource":"-1067234240","Value":"0","scenarioTime":"06:46"}
06:46        GL_Power_Wasted                       {"Resource":"-665414129","Value":"0","scenarioTime":"06:46"}
06:46        GL_Wind_Power_Produced                {"Resource":"-626004793","Value":"0","scenarioTime":"06:46"}
06:46        GL_Coal_Power_Produced                {"Resource":"1467018548","Value":"34650","scenarioTime":"06:46"}
06:47        GL_Action_ToolCategory  {"action":"opened","tool":"power","scenarioTime":"06:47"}
06:46        GL_Air_Pollution          {"Resource":"295846734","Value":"43135844","scenarioTime":"06:46"}
07:00        GL_Unit_Plop                          {"UGuid":"0x9122c84d","name":"","Pos":"-237.33, 233.38, 146.93","scenarioTime":"07:00"}
07:01        GL_Dezone                 {"type":"commercial","scenarioTime":"07:01"}
07:02        GL_Action_Building        {"action":"selected","name":"Solar Power Plant","scenarioTime":"07:02"}
07:02        GL_Action_Building        {"action":"viewed","name":"Solar  Power Plant","scenarioTime":"07:02"}
07:03        GL_Action_Building        {"action":"deselected","name":"Solar Power Plant","scenarioTime":"07:03"}
07:03        GL_Action_ToolCategory  {"action":"closed","tool":"power","scenarioTime":"07:03"}
07:04        GL_Action_Building        {"action":"view-hidden","name":"Solar  Power Plant","scenarioTime":"07:04"}
07:08        GL_Unit_Plop              {"UGuid":"0xa230f2dc","name":"","Pos":"-147.38, 327.56, 146.93","scenarioTime":"07:08"}
07:15        GL_Challenge_Heartbeat  {"jobs":"6062","name":"Medusa A3 - Large City.txt","pollution":"86402071","simoleons":"9310","scenarioTime":"07:15"}
```

Insights in the Development of Psychometrics / Educational Measurement

- Probability-based reasoning, for managing evidence.
- Building models that suited an inferential problem cast in some psychological theory, with pertinent data.
- Seeing reliability, validity, comparability, generalizability, and fairness not just as measurement issues, but "social values that have meaning and force outside of measurement wherever evaluative judgments and decisions are made."

(Messick, 1994)

The Standard Ed Measurement Paradigm

Psychology

- Ed measurement paradigm: observation & control (150 years) is a layer over the Examination paradigm (2000 years!)
- Not much focus on cognitive or learning processes.

Data (key role for data mining and learning analytics)

- Human ratings of performances hide complexity, & don't scale.
- "Objective scoring" does scale and can be automated, but at cost of constraining observational situations and performances.

Models

- Galton, Cattell, Spearman, Thurstone, etc. were tackling problems jointly in psychology, observation methods, modeling, and statistics.
- Early learning analytics / data mining: Regression, correlation, multidimensional scaling, cluster analysis, factor analysis, path diagrams.

The Standard Ed Measurement Paradigm

Probability-Based Reasoning

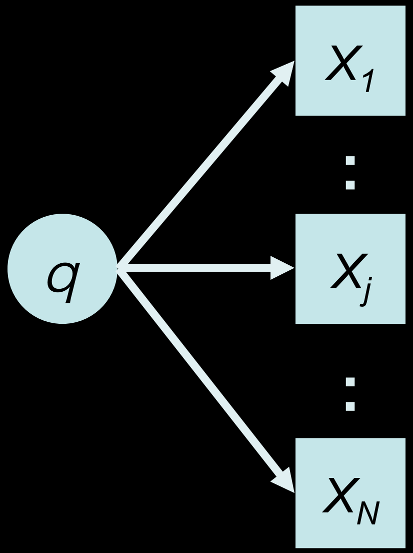
Probability isn’t really about numbers;
it’s about the structure of reasoning.

Glenn Shafer (quoted in Pearl, 1988)

The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Classical Test Theory



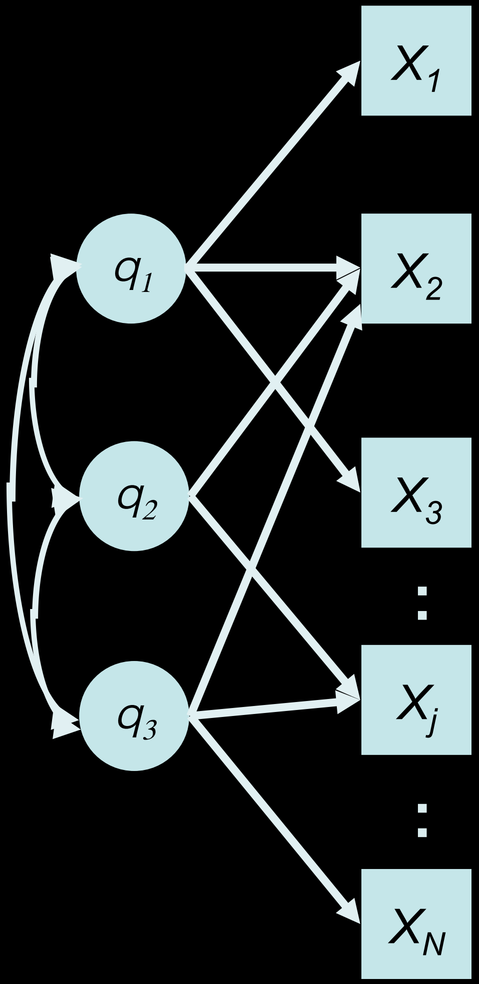
Conditional independence posited among *X*s given *q*.

Item response theory (IRT) has same structure at the level of items rather than tests.

The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Factor analysis



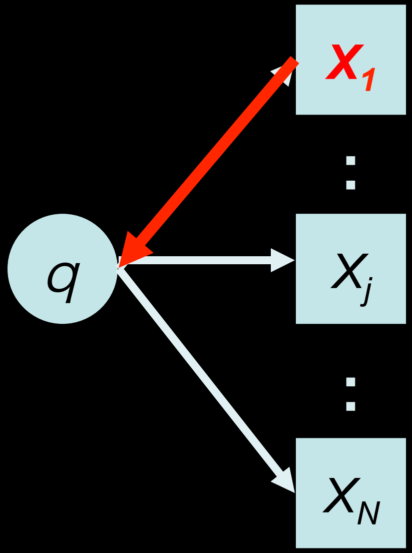
Both discovery and guided exploration of underlying, psychologically-relevant, structure to “explain” patterns in data.

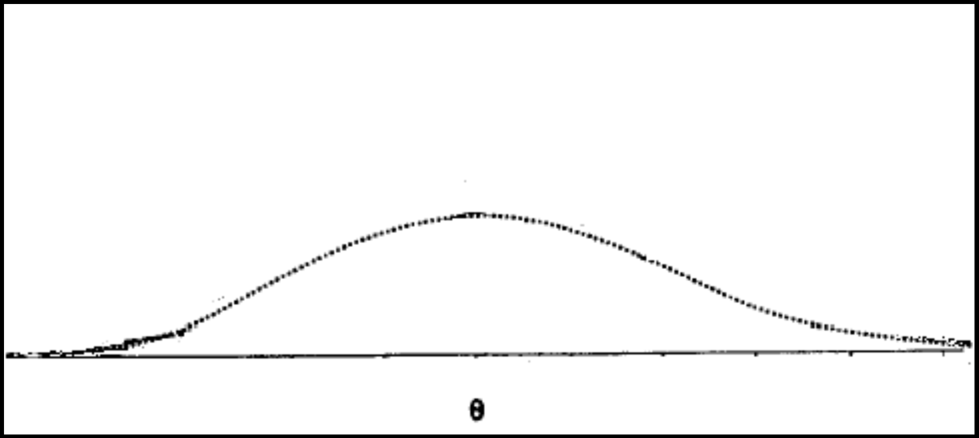
Same basic idea as current exploratory use of Bayes nets, multidimensional scaling, Gaussian mixture cluster analysis.

The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Bayesian inference

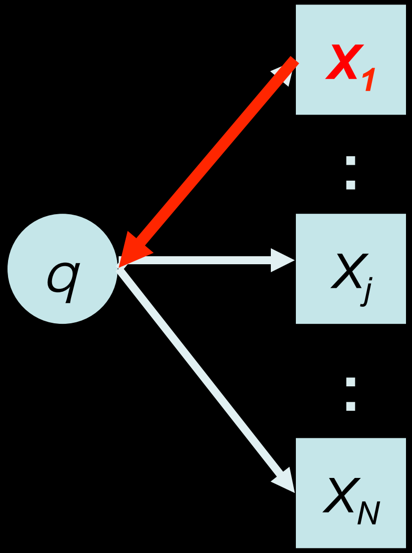




The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Bayesian inference

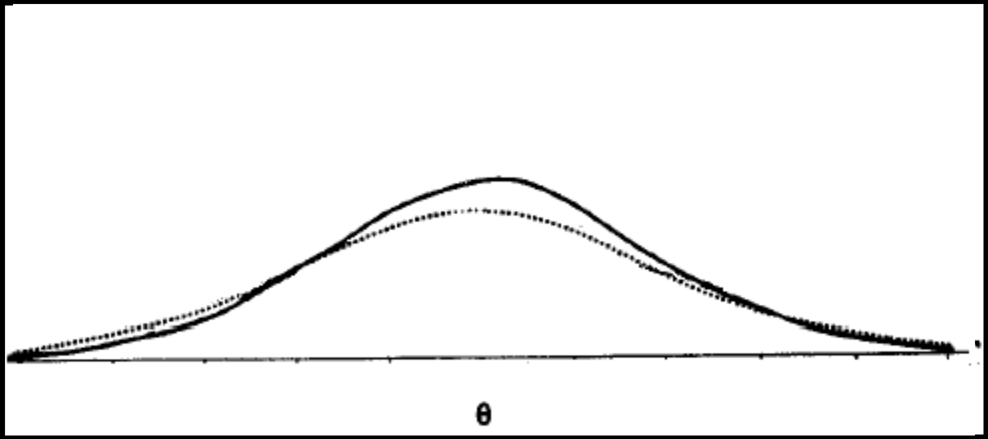


X_1

q

X_j

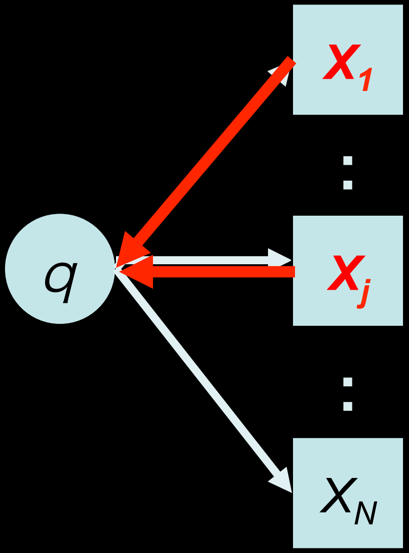
X_N

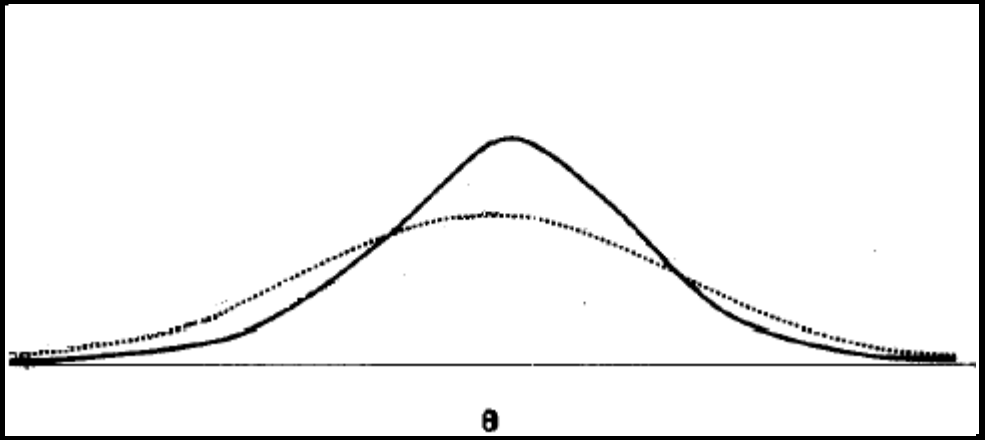


The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Bayesian inference

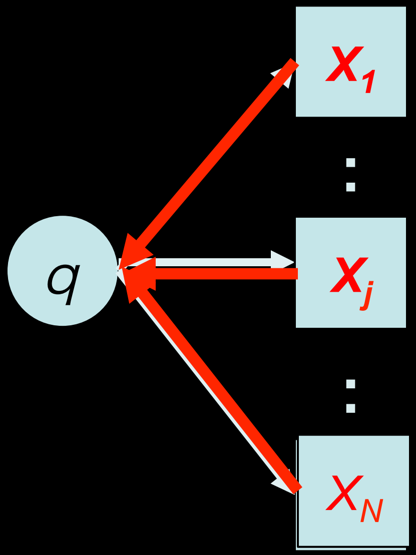


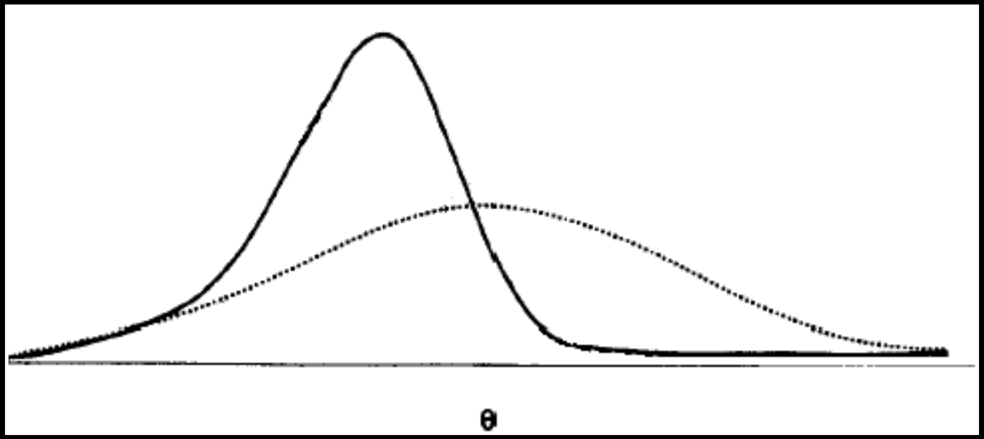


The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Bayesian inference

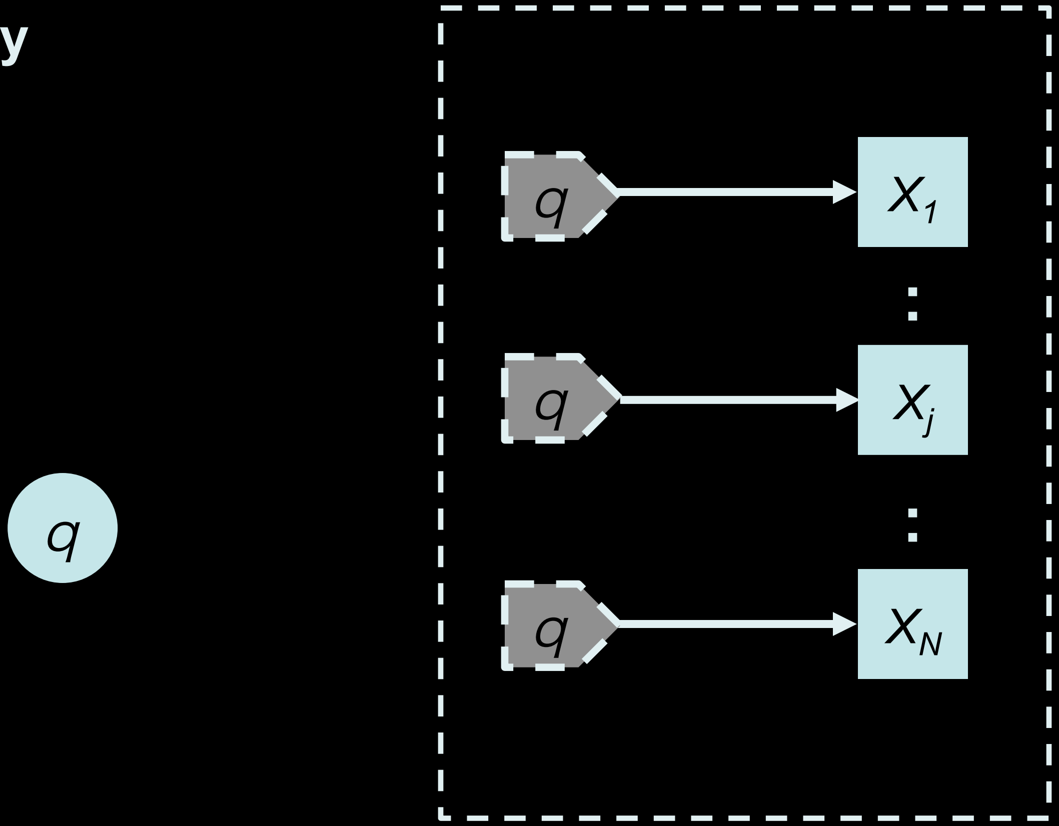




The Standard Ed Measurement Paradigm

Probability-Based Reasoning

- Modularity



The Standard Ed Measurement Paradigm

Probability-Based Reasoning

- Metric for quantifying evidence.
- Common framework for synthesizing different observations for different people.
- Tools to investigate how well do the patterns the model can express accord with the patterns that are in the data.

This conceptual framework, and these practical advantages, can extend to inference to assessments richer than SEMP in several ways.

The Standard Ed Measurement Paradigm

Social Values

Validity, reliability, comparability, [generalizability,] and fairness are not just measurement issues, but *social values* that have meaning and force outside of measurement wherever evaluative judgments and decisions are made.

Messick, 1994

Situative / Sociocognitive Psychology

Most approaches to curriculum, instruction, and assessment are based on theories and models that have not kept pace with modern knowledge of how people learn.

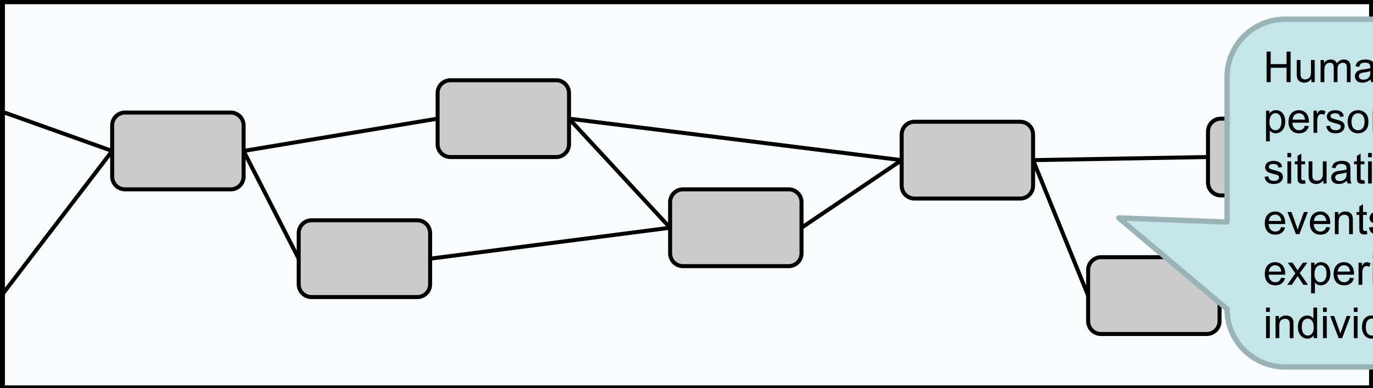
They are based on implicit and limited conceptions of learning that tend to be fragmented, outdated, and poorly delineated for subject-matter domains.

Jim Pellegrino (2016)

Situative / Sociocognitive Psychology

Confluence of ideas & research across domains –

- e.g., learning sciences; domain-based learning; sociolinguistics; “new literacy”; anthropology; cognitive, situated, social, neuro psychology.

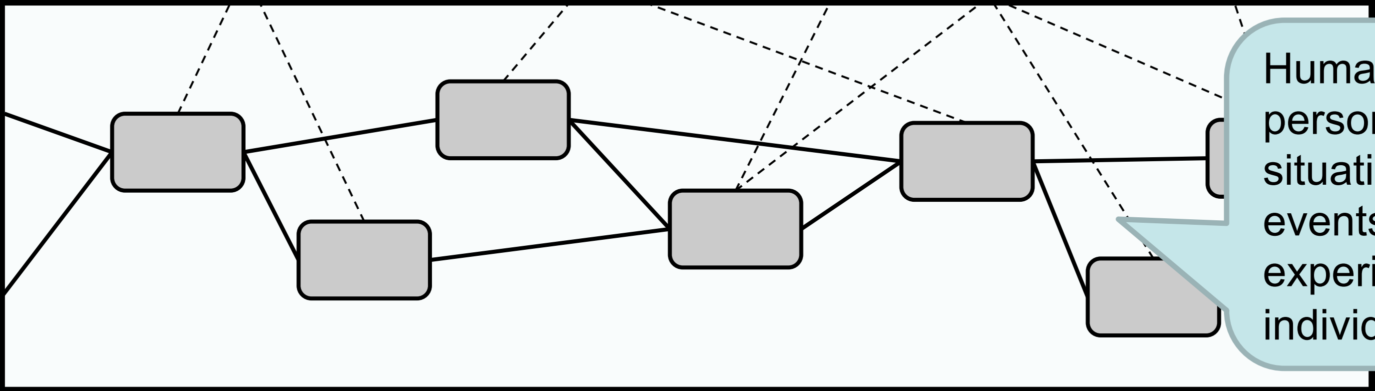


Human-level activity, persons acting within situations--the actions, events, and activities we experience as individuals.

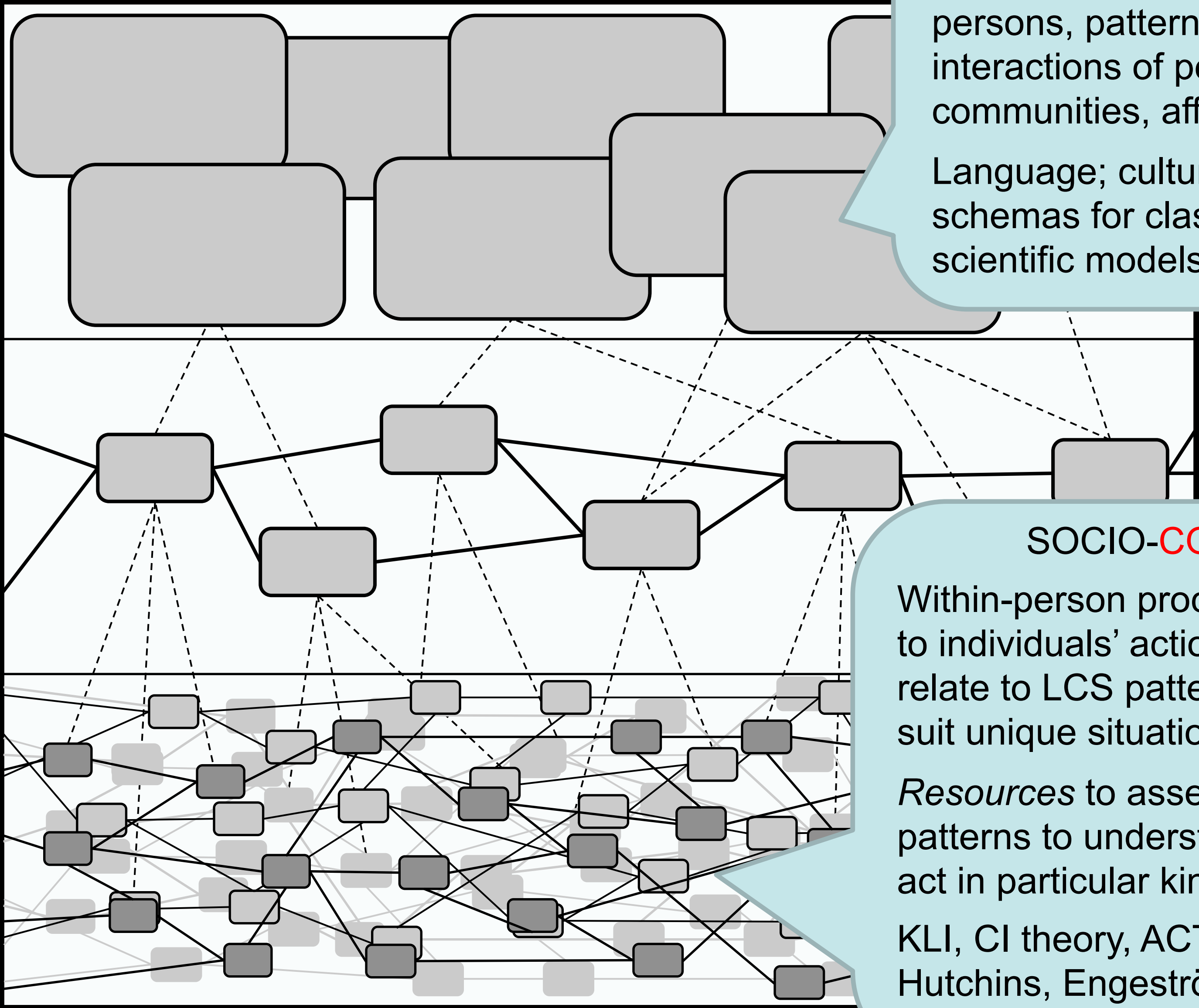
SOCIO-COGNITIVE

Extrapersonal, or between-persons, patterns: Regularities in interactions of people in communities, affinity spaces.

Language; cultural models; schemas for classrooms; scientific models. (LCS patterns)



Human-level activity, persons acting within situations--the actions, events, and activities we experience as individuals.



SOCIO-COGNITIVE

Extrapersonal, or between-persons, patterns: Regularities in interactions of people in communities, affinity spaces.

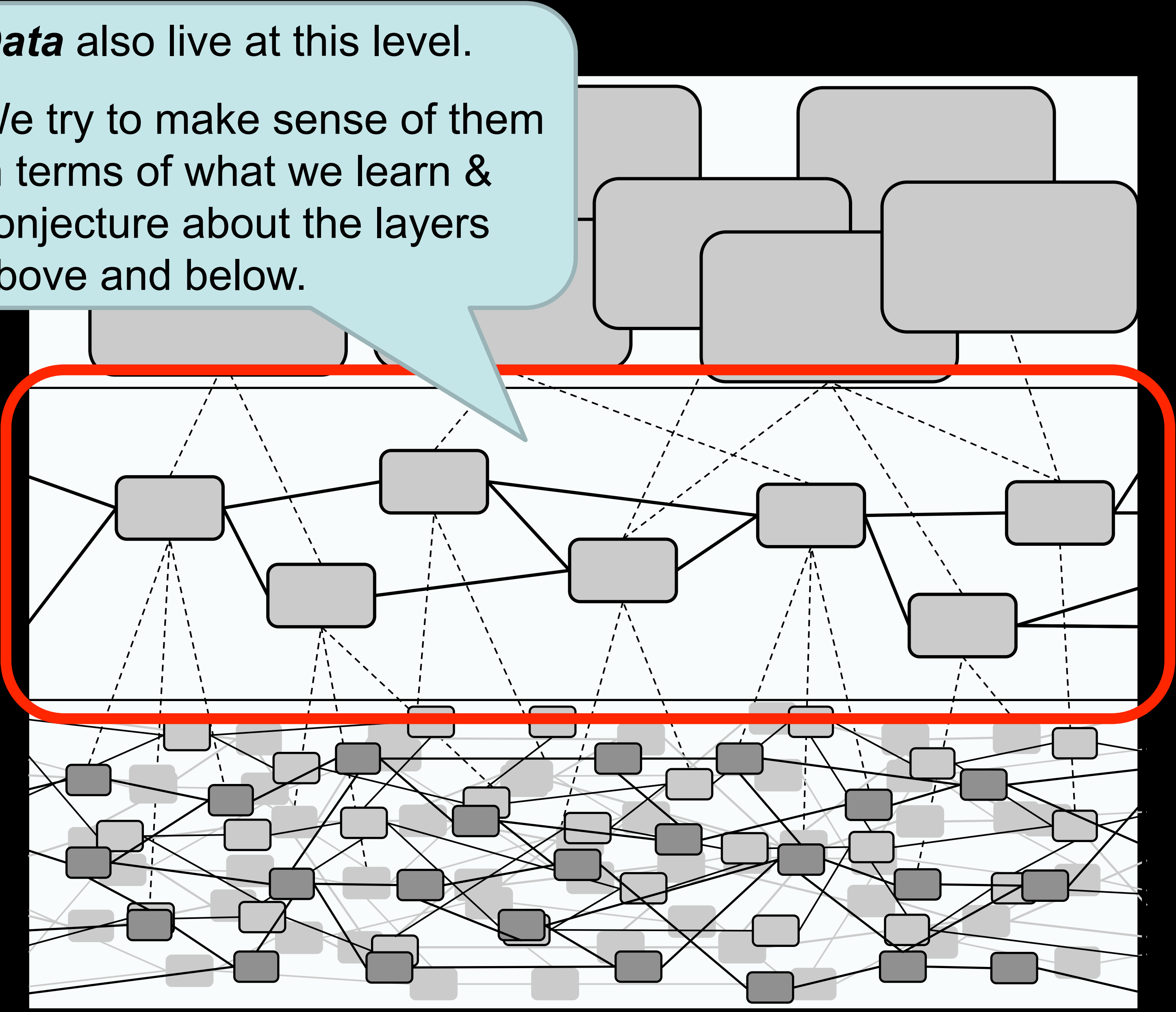
Language; cultural models; schemas for classrooms; scientific models. (LCS patterns)

SOCIO-COGNITIVE

Within-person processes give rise to individuals' actions. Must both relate to LCS patterns and adapt to suit unique situations.

Resources to assemble particular patterns to understand, create, & act in particular kinds of situations.

KLI, CI theory, ACT-R; Lave, Hutchins, Engeström; Language as a complex adaptive system.



Data also live at this level.

We try to make sense of them in terms of what we learn & conjecture about the layers above and below.

Situative / Sociocognitive Psychology

Person acting in situation.

- What is important to notice?
- What does it mean?
- What will happen next?
- What kinds of things can I say / do next?
- How can I create / negotiate situations?

What does this imply for assessment?

- A great change in psychology and implied task environments… which changes what the variables and distributions mean.

Implications for Psychometric Models

Q: How do we think of constructs (hence, latent variables)?

A: Tendencies / capabilities / manners of perceiving, processing, and acting in certain kinds of situations—constellations of certain *kinds* of resources.

But thinking in terms of resources that are…

- Idiosyncratic, but similarities due to practices and LCS patterns that structure situations.
- Contingent, and local in time and associations among people.
- Initially strongly connected to contexts of learning.

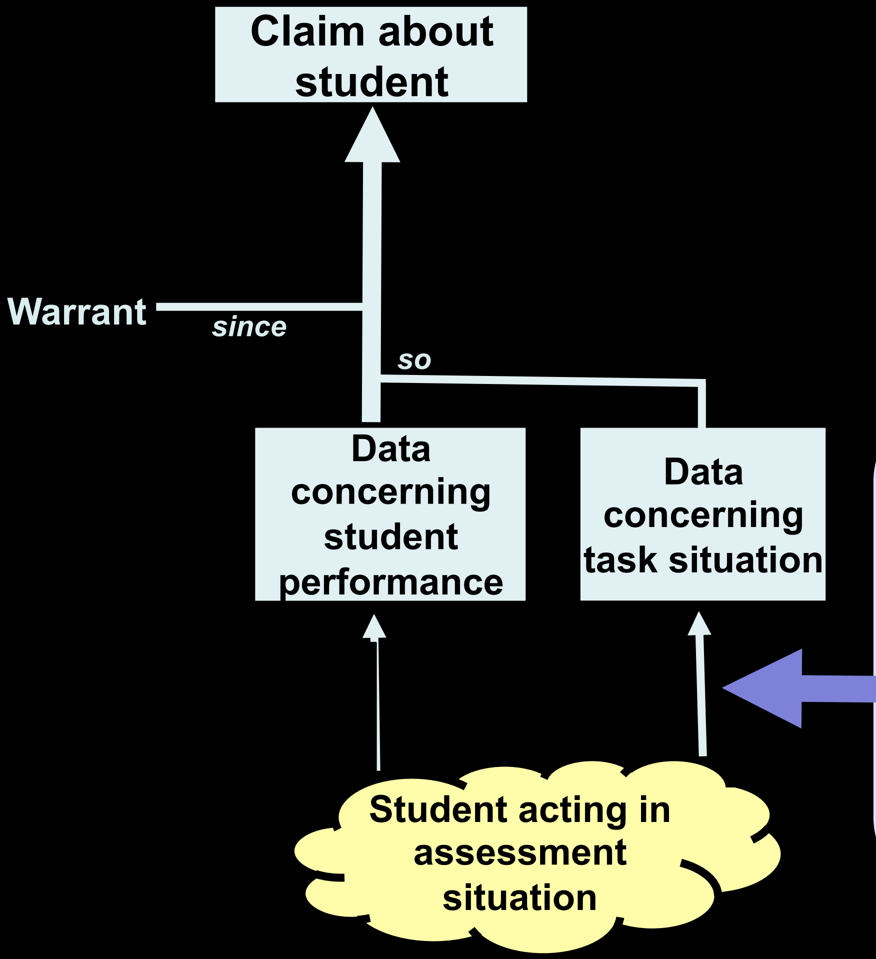
What is the range of a model's "as if" usefulness?

For what purposes?

Implications for Environments & Models

- Continuous activity.
  - We must characterize evidence, not “score responses.”
- Examinee actions change the situation.
- Changing proficiencies (esp. learning).
- Multiple proficiencies.
- Conditional dependence.
- Different proficiency / observable combinations.
- Multiple modalities.
- Interaction among examinees (e.g., collaboration).

The structure of assessment arguments



Claim about student

Warrant

since

so

Data concerning student performance

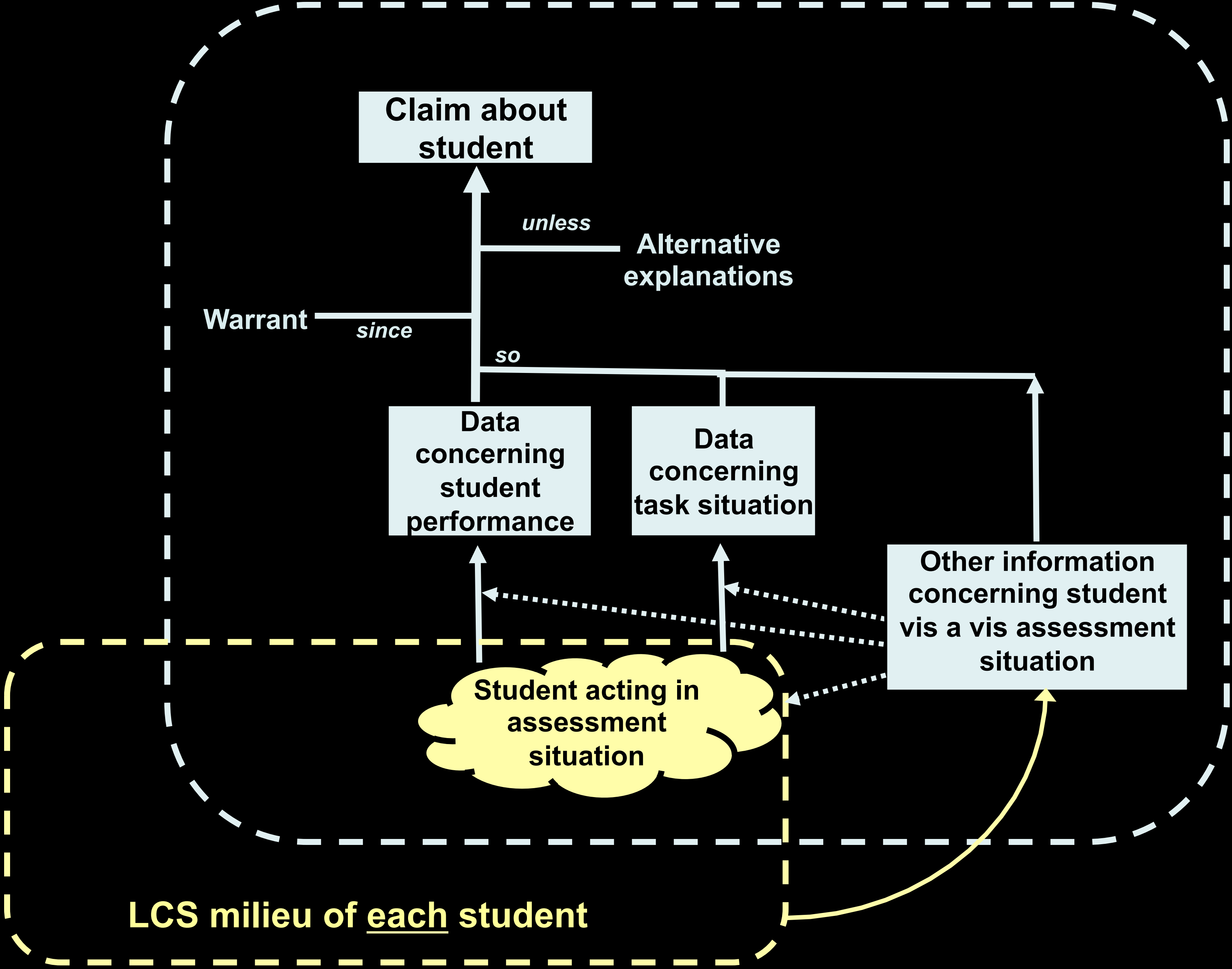
Data concerning task situation

Student acting in assessment situation

Concerns features of (possibly evolving) context as seen from the view of the assessor – in particular, those seen as relevant to targets of inference.

More on this too.

Social / cultural contextualization of assessment



Claim about student

unless

Alternative explanations

Warrant

since

so

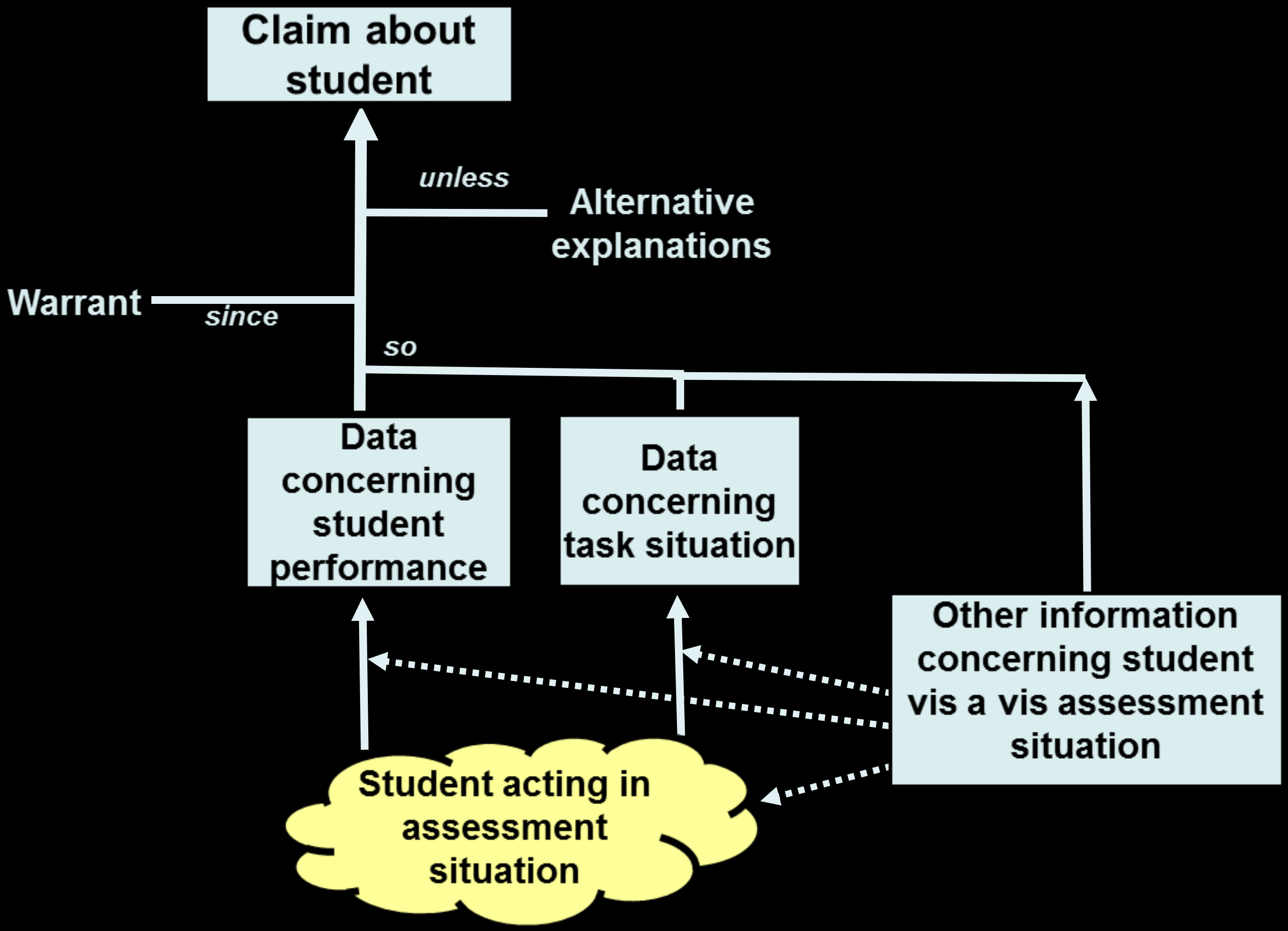
Data concerning student performance

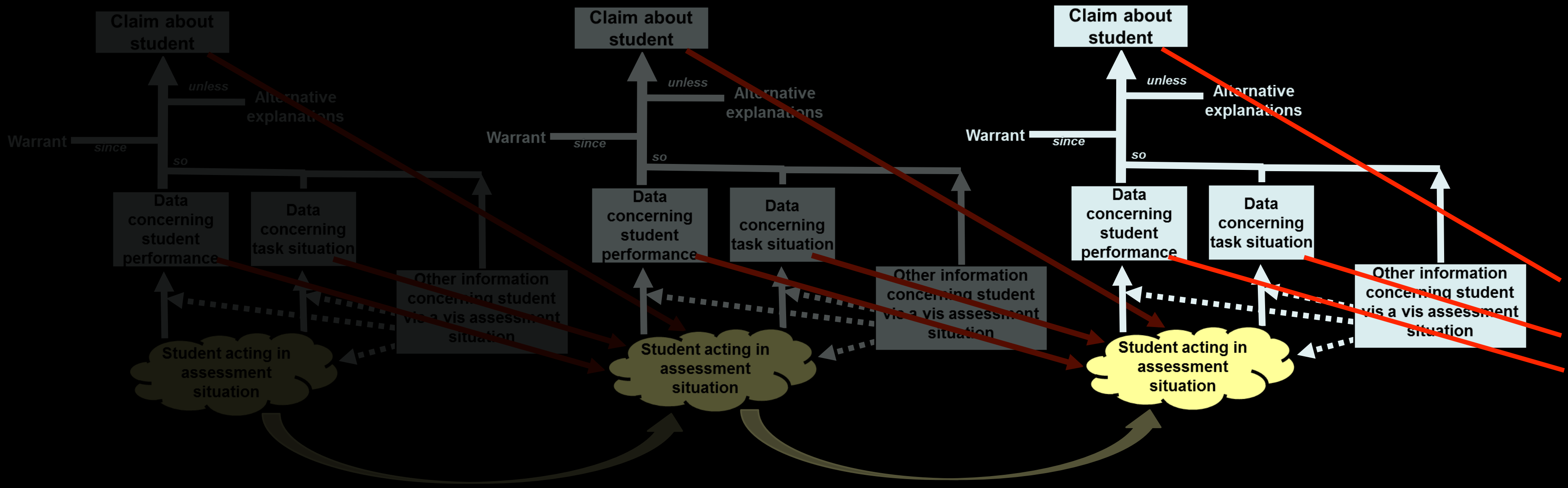
Data concerning task situation

Other information concerning student vis a vis assessment situation

Student acting in assessment situation

LCS milieu of each student


Claim about student
unless
Alternative explanations
Warrant
since
so
Data concerning student performance
Data concerning task situation
Other information concerning student vis a vis assessment situation
Student acting in assessment situation



Claim about student

unless

Alternative explanations

Warrant

since

so

Data concerning student performance

Data concerning task situation

Other information concerning student vis a vis assessment situation

Student acting in assessment situation

Claim about student

unless

Alternative explanations

Warrant

since

so

Data concerning student performance

Data concerning task situation

Other information concerning student vis a vis assessment situation

Student acting in assessment situation

Claim about student

unless

Alternative explanations

Warrant

since

so

Data concerning student performance

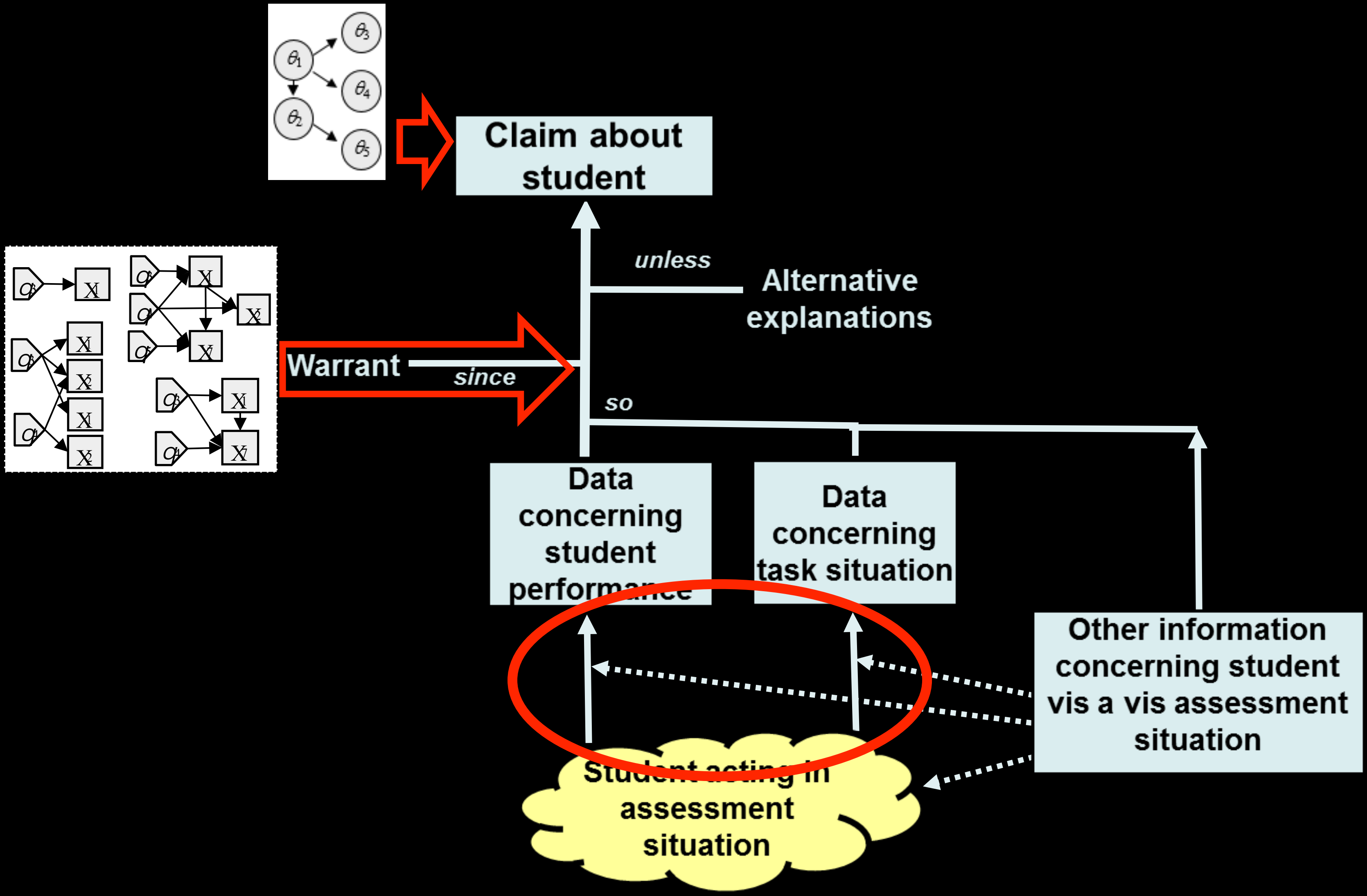
Data concerning task situation

Other information concerning student vis a vis assessment situation

Student acting in assessment situation

–

Instantiating an assessment argument in objects and processes



Claim about student

unless

Alternative explanations

Warrant

since

so

Data concerning student performance

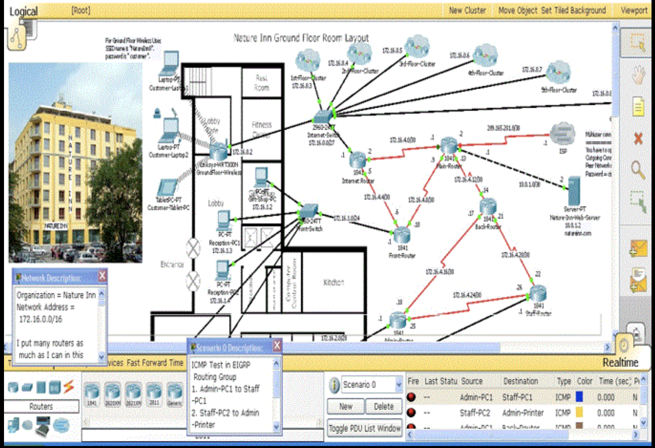
Data concerning task situation

Other information concerning student vis a vis assessment situation

Student acting in assessment situation

New Forms of Assessment

SimCityEDU. GlassLab



Network Description:

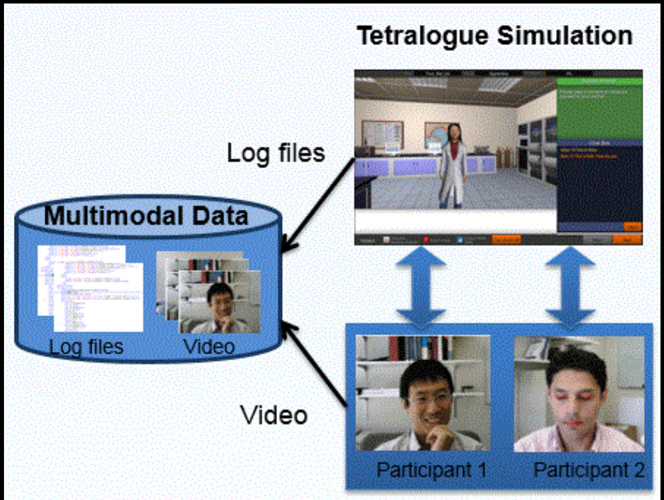
Organization = Nature Inn
Network Address =
172.16.0.0/16

I put many routers as
much as I can in this

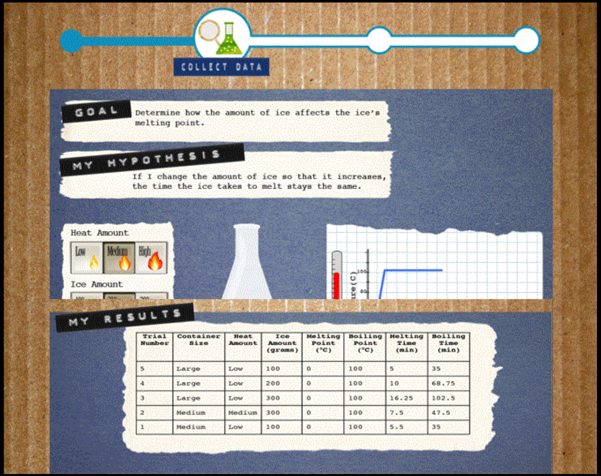
Scenario 0 Description:

ICMP Test in EIGRP
Routing Group
1. Admin-PC1 to Staff
-PC1
2. Staff-PC2 to Admin
-Printer

Packet Tracer. Cisco Networking Academy
Behrens & DiCerbo, 2013



Tetralogues. Khan & Suendermann-Oeft, ETS



GOAL Determine how the amount of ice affects the ice's melting point.

MY HYPOTHESIS If I change the amount of ice so that it increases, the time the ice takes to melt stays the same.

MY RESULTS

Trial Number	Container Size	Heat Amount	Ice Amount (grams)	Melting Point (°C)	Boiling Point (°C)	Melting Time (min)	Boiling Time (min)
5	Large	Low	100	0	100	5	35
4	Large	Low	200	0	100	10	68.75
3	Large	Low	300	0	100	16.25	102.5
2	Medium	Medium	300	0	100	7.5	47.5
1	Medium	Low	100	0	100	5.5	35

Sao Pedro, Gobert, Toto, & Paquette, AERA 2015

New Forms of Assessment

Constructs

- Systems thinking, Interactional speaking, Troubleshooting, Cross-cultural communication, Inquiry, Collaboration.

Activity Models (née Task Models)

- Simulation spaces, Trialogue w avatars, Inquiry space.
  Situations & interactions designed to evoke evidence.

Work Product(s)

- Log files, videos, artifacts, speech/chats, artifacts/designs.

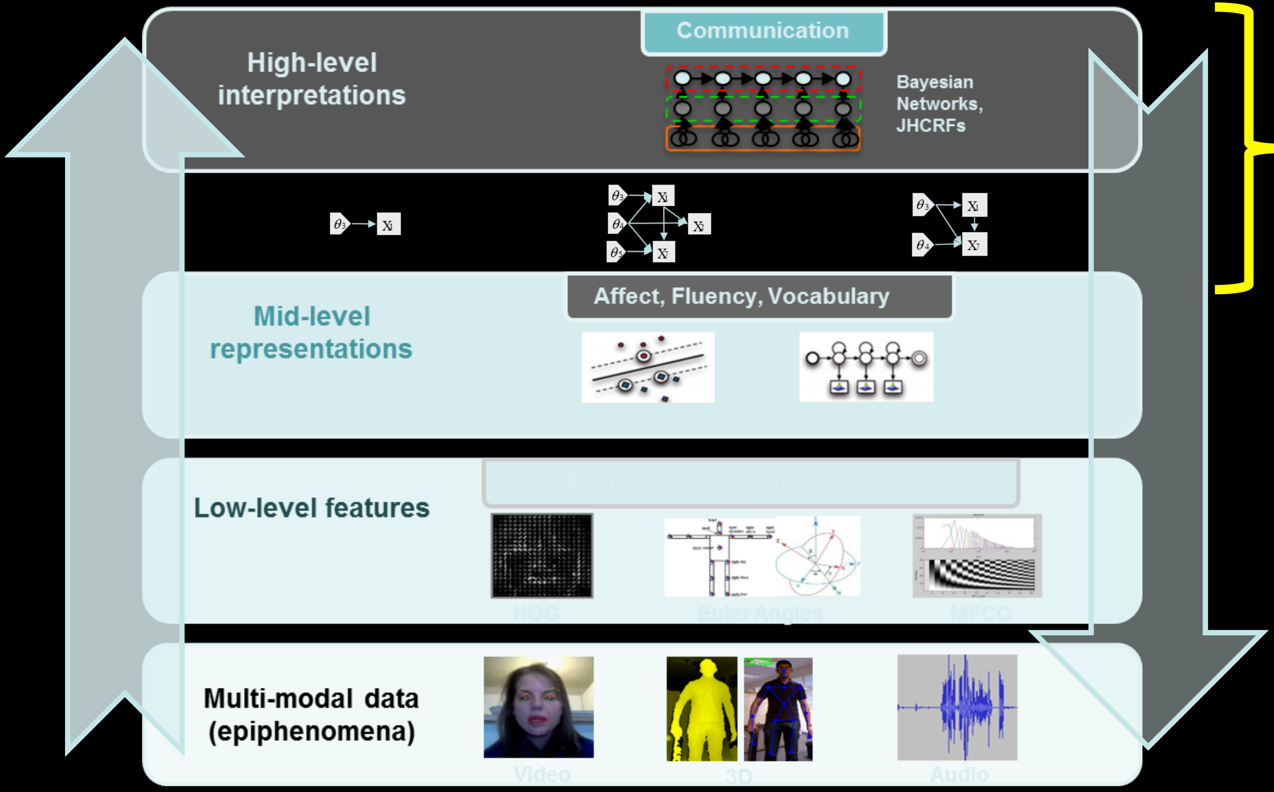
Psychometric Models

- SMVs tuned to theory, data, interaction, & purpose.
  OVs allow different particulars same construct-driven theory.

Evidence Identification…

“Computational Psychometrics”

Evidence for constructs from low-level data.
Hierarchies of chain of evidentiary reasoning (can be up & down, theory-aided.)


High-level interpretations
Communication
Bayesian Networks, JHCRFs
Mid-level representations
Affect, Fluency, Vocabulary
Low-level features
HOG
Euler Angles
MFCC
Multi-modal data (epiphenomena)
Video
3D
Audio

Psychometric models

Machine Learning,

Theory,

Data mining

Theory

Design

Supervised,
Semisupervised,
Unsupervised

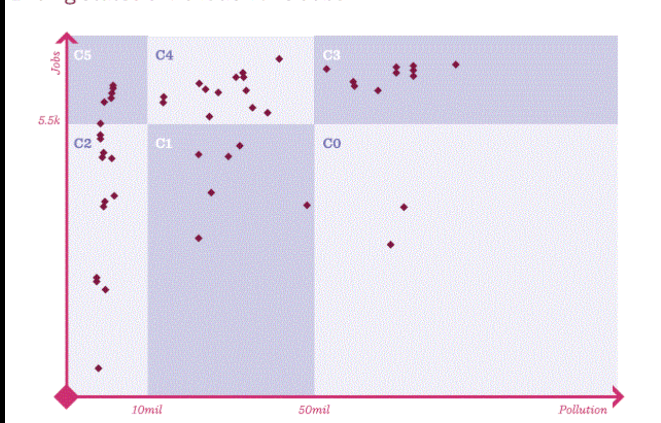
from von Davier, Khan, & Kerr

Hierarchical Inference in Evidence Identification

Khan & Kerr (2014, 2015)

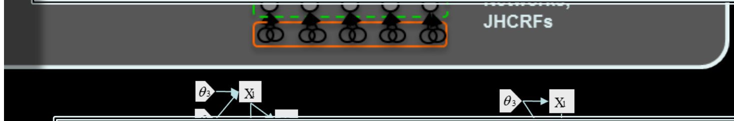
SimCityEDU: Pollution Challenge!

Ending States of Pollution and Jobs



Level	Cell	Evaluation Rule
5	C5	Pollution < 10mil Jobs > 5.5k
4	C4	Pollution > 10mil and < 50mil Jobs > 5.5k
3	C3	Pollution > 50mil Jobs > 5.5k
3	C2	Pollution < 10mil Jobs < 5.5k
2	C1	Pollution > 10mil and < 50mil Jobs < 5.5k
1	C0	Pollution > 50mil Jobs < 5.5k

Construct was levels on a systems-thinking learning progression variable – reflects *kinds* of things people can do in *kinds* of situations. Model incorporated *q* change at the level of challenges.



Psychometric models

Summary functions of counts of these actions and system-state variables are input variables into a dynamic Bayes net – hidden Markov model with respect to level on learning progression.

high-pollution one.

Locations, times, durations and objects of "verb clauses"– verbs like "rezone," "bulldoze," "query map." → Log file contents. (+ system actions)

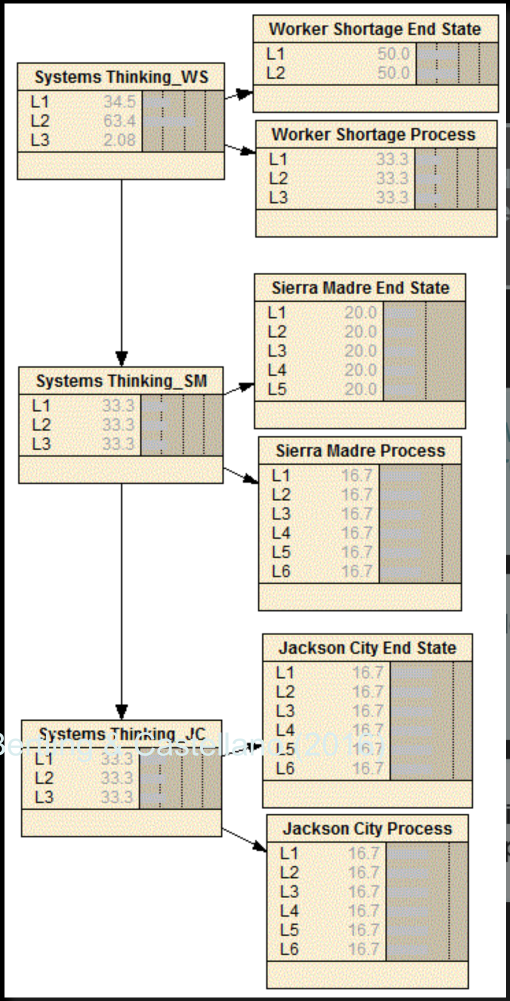
Multi-modal data (epiphenomena)

Locations, times, and durations of clicks, hovers, drag & drops, etc.

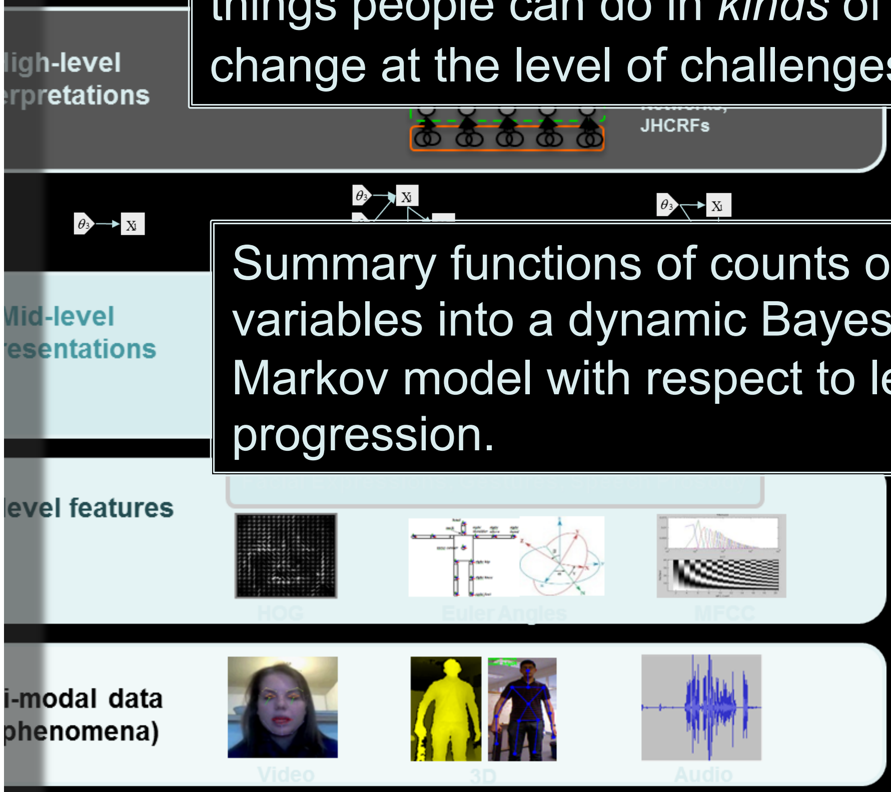
Hierarchical Inference in Evidence Identification

Khan & Kerr (2014, 2015)

SimCityEDU



Bertling & Castellano (2016)



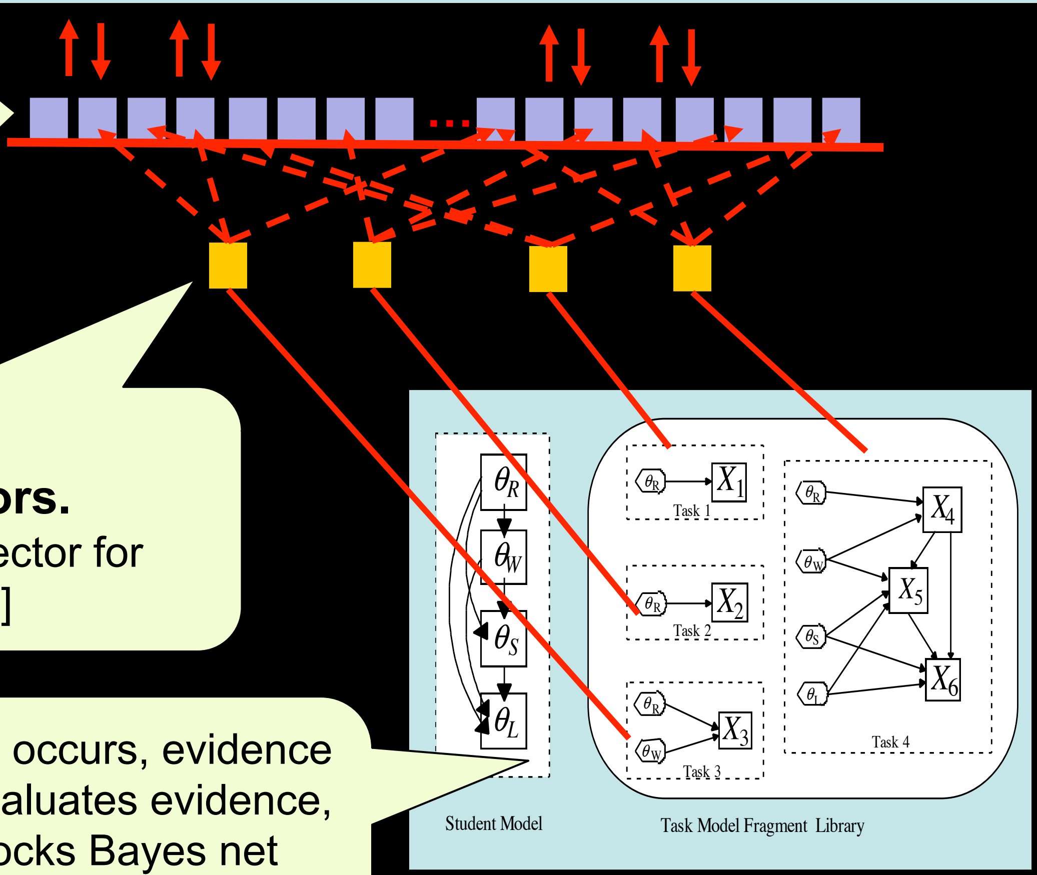
Construct is levels on a systems-thinking learning progression variable – reflects *kinds* of things people can do in *kinds* of situations. Model change at the level of challenges.

Psychometric models

Summary functions of counts of these are input variables into a dynamic Bayes net – hidden Markov model with respect to level on learning progression.

Flow of Activity

State vector. Tracks relevant features of situations and past actions.



Psychometric objects and processes

Evidence-bearing opportunity detectors. Agents monitor state vector for EBOs. [beyond "tasks"]

When a particular EBO occurs, evidence identification routine evaluates evidence, and "scoring engine" docks Bayes net fragment with proficiency model to update probability distribution for *qs*.

Social values, revisited

Validity, reliability, comparability, [generalizability], and fairness are not just measurement issues, but *social values* that have meaning and force outside of measurement wherever evaluative judgments and decisions are made.

Messick, 1994

Conclusion – Key Ideas

- Probability-based reasoning.
  - Manage evidence
  - Address reliability, validity, generalizability, comparability, fairness
- Situative / Sociocognitive psychological perspective.
- “Assessment as measurement”

Conclusion – Key Ideas

- Dialectic between design and discovery.

- “Computational psychometrics”: Synergy of psychometrics, learning analytics, data mining.

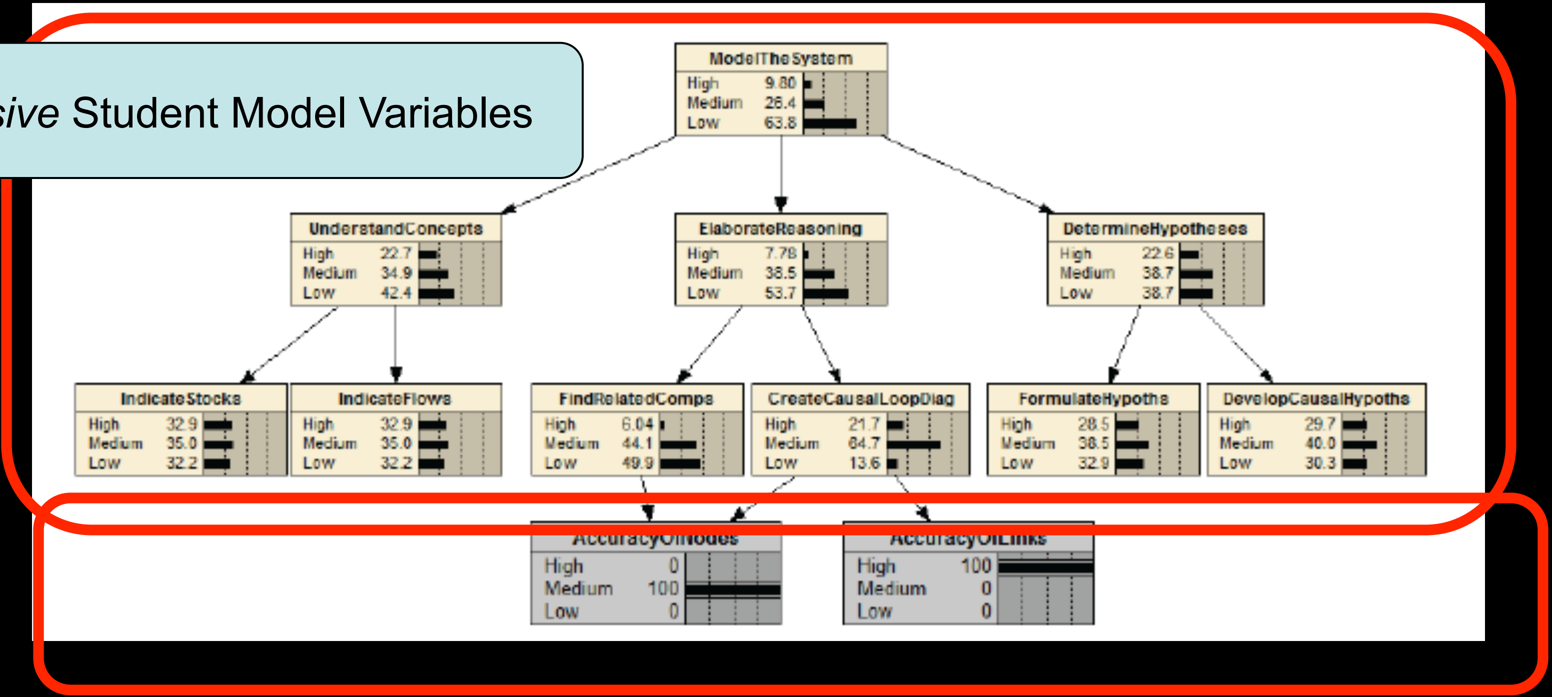
- Validity, reliability, comparability, generalizability, fairness
  - Probability models help address them rigorously.

Thank you.

A Couple Quick Examples

Modular Bayes net for Evaluating a Casual-Loop Diagram

Pervasive Student Model Variables

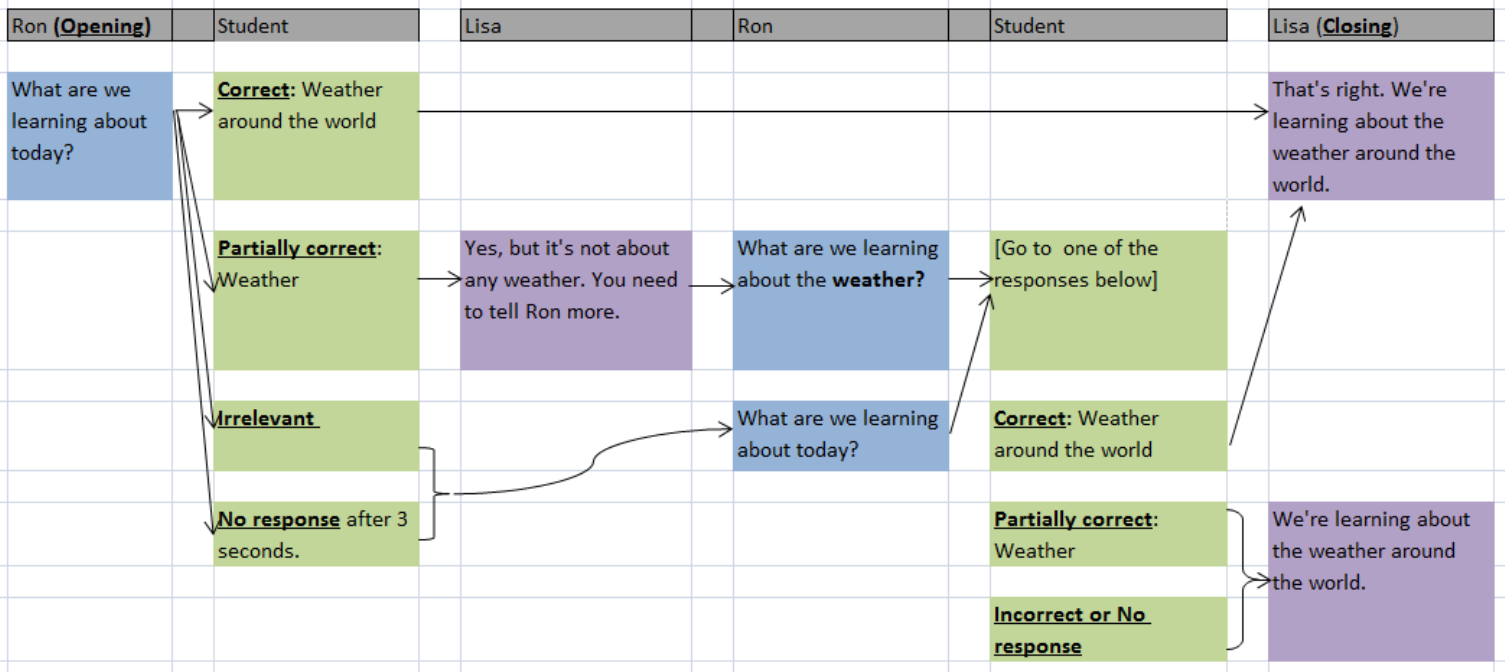


Ephemeral Observable variables from an evidence-bearing opportunity

Shute et al. (2010)

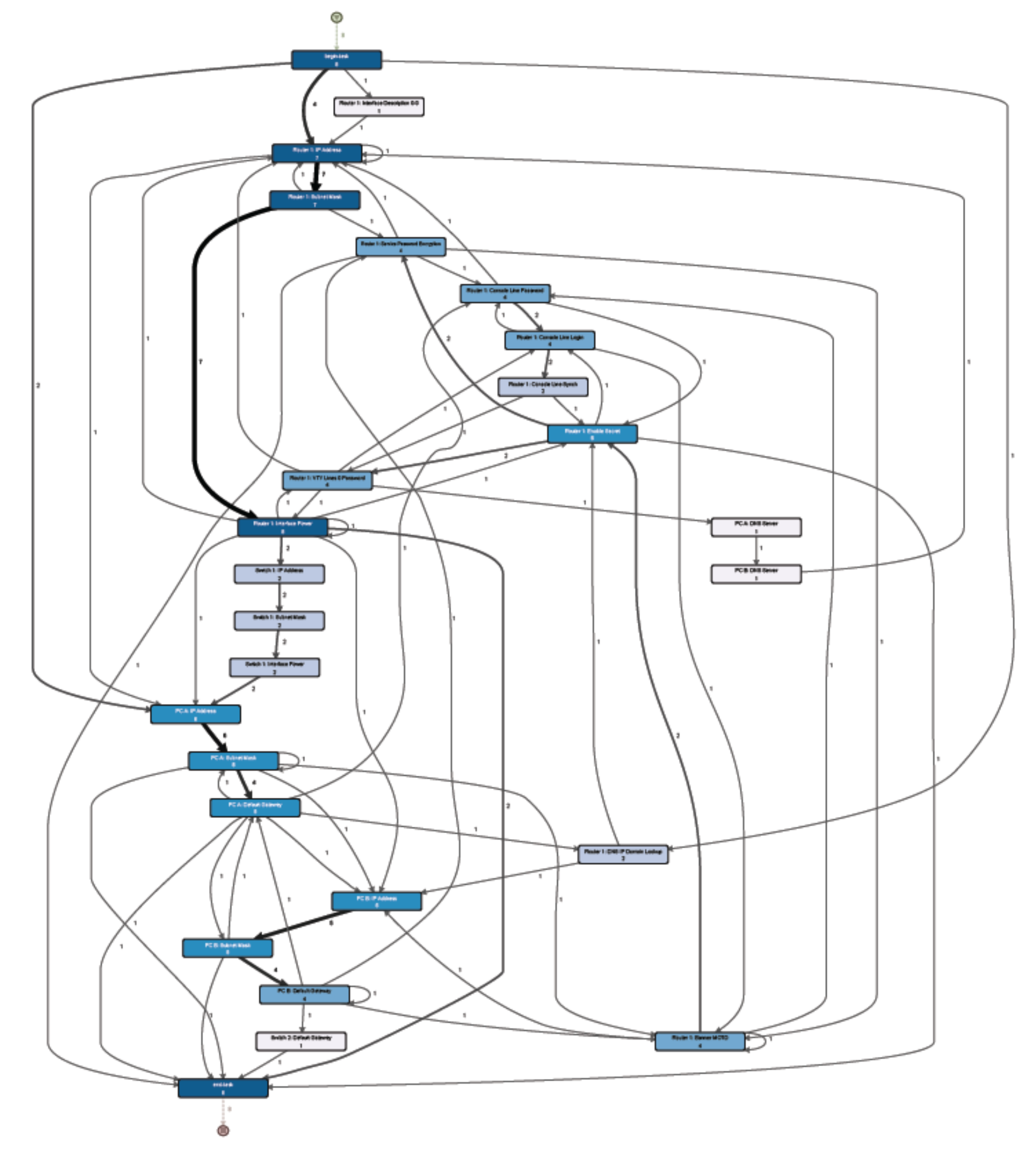
Conversation Mapping in Trialogue Assessment

Framework for using NLP with chat with avatars, to monitor and CREATE evidence-bearing opportunities.



Ron (**Opening**)	Student	Lisa	Ron	Student	Lisa (**Closing**)
What are we learning about today?	**Correct**: Weather around the world				That's right. We're learning about the weather around the world.
	Partially correct: Weather	Yes, but it's not about any weather. You need to tell Ron more.	What are we learning about the **weather?**	[Go to one of the responses below]	
	Irrelevant		What are we learning about today?	**Correct**: Weather around the world	
	No response after 3 seconds.			**Partially correct**: Weather	We're learning about the weather around the world.
				Incorrect or No response	

LaMar & Bergner (2015)



Business-Process Modeling to Identify Computer-Network Troubleshooting Patterns of Experts and Novices

Cisco Networking Academy's Packet Tracer tasks.

Tiago Calico (2016)

Using network theory to improve task design and scoring

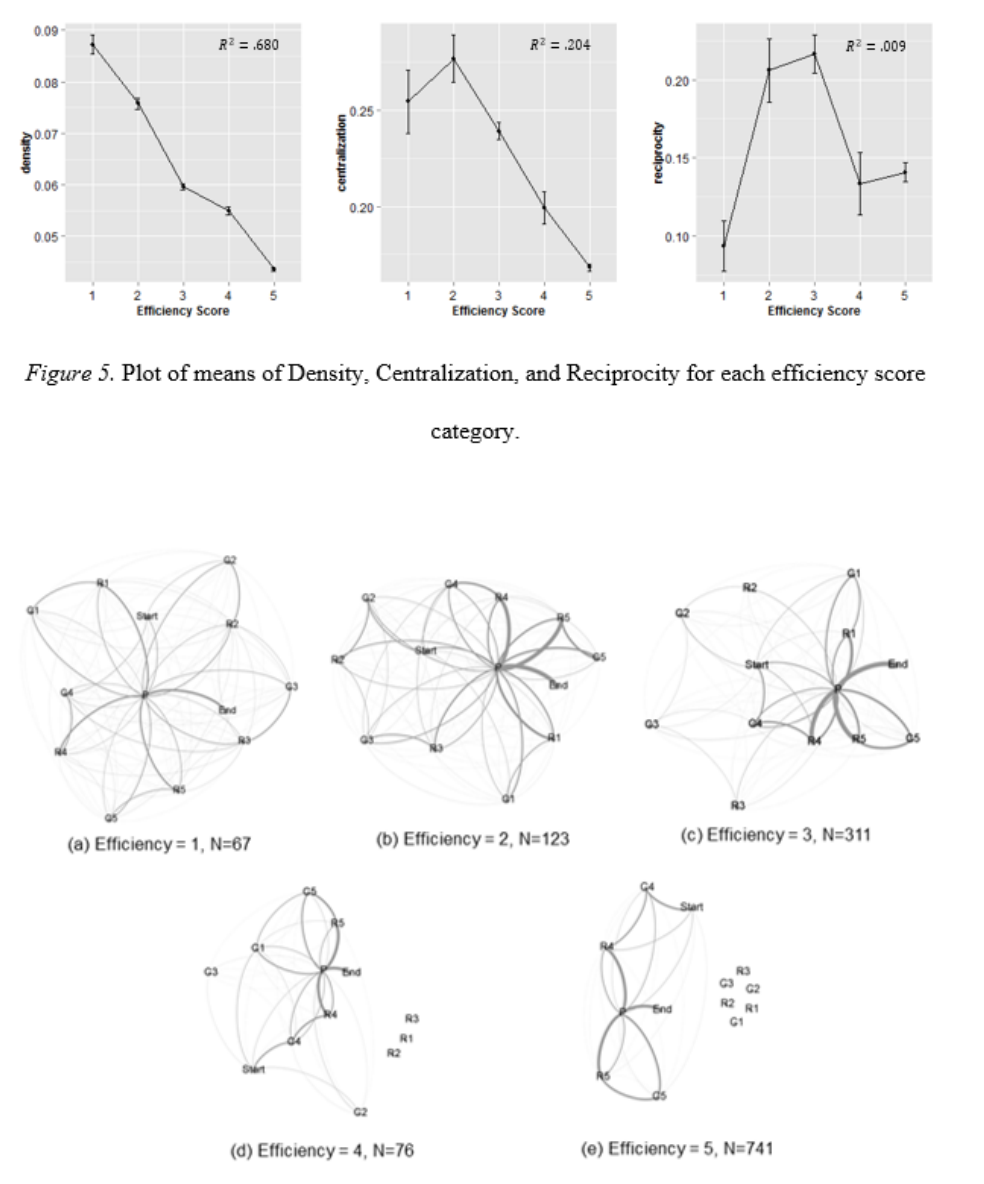


Figure 5. Plot of means of Density, Centralization, and Reciprocity for each efficiency score category.

(a) Efficiency = 1, N=67

(b) Efficiency = 2, N=123

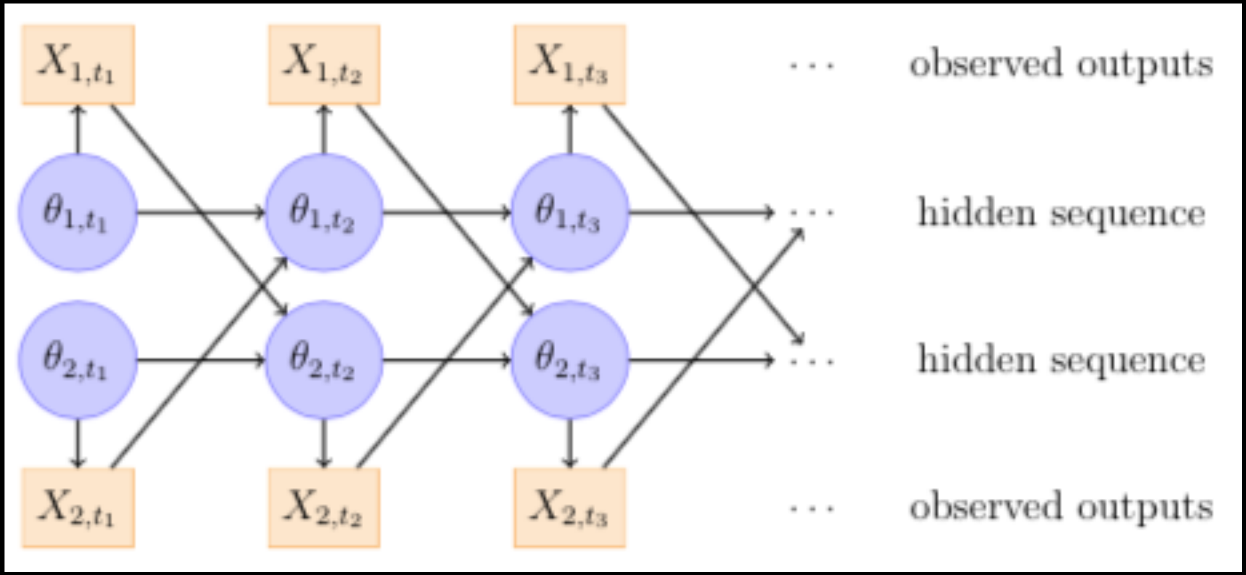
(c) Efficiency = 3, N=311

(d) Efficiency = 4, N=76

(e) Efficiency = 5, N=741

Zhu, Shu, & von Davier (2016)

A Hidden Markov Model for Collaboration



LaMar & Bergner (2015)

The Standard Ed Measurement Paradigm

Probability-Based Reasoning

Ba