

A Sociocognitive Perspective on Educational Measurement



Keynote A

inferences

A Sociocognitive Perspective on Educational Measurement

Robert J. Mislevy

Emeritus Professor
University of Maryland at College Park

Keynote Address, *Beyond results 2021: From log data to valid inferences*
IEA / DIPF / ZIB, September 30, 2021

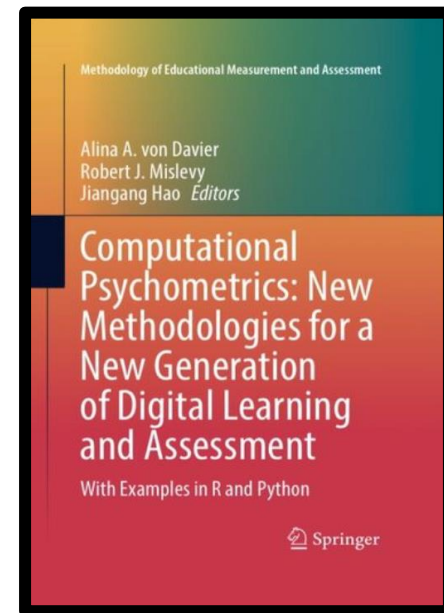


Foundations

- Sociocognitive psychological perspective
- Assessment-as-Argument Structuring



Mislevy, R.J. (2018) *Sociocognitive foundations of educational measurement*. NY/London: Routledge.

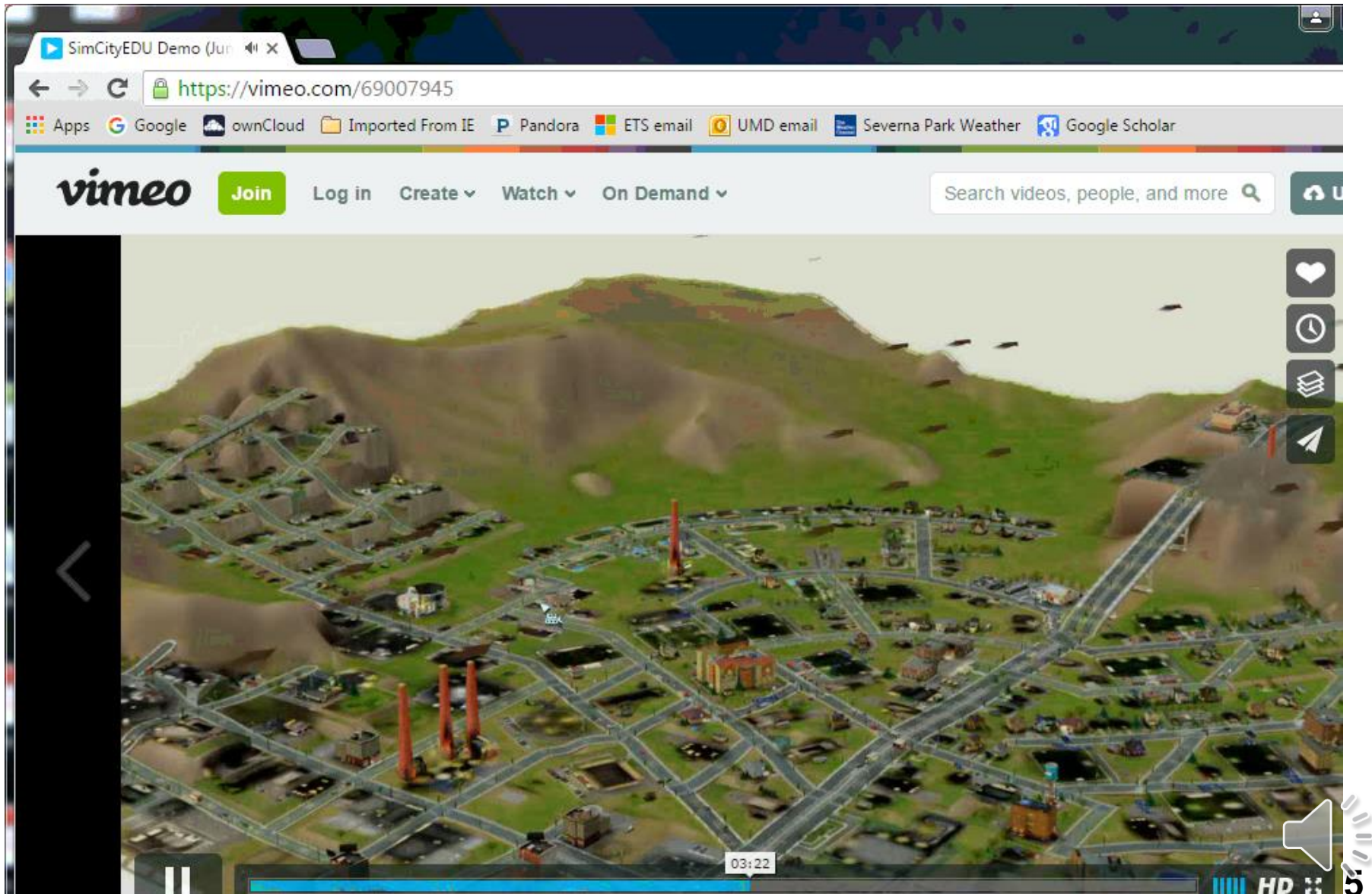


Von Davier, A., Mislevy, R.J., & Hao, J. (in press). *Computational psychometrics: New methodologies for a new generation of digital learning and assessment*.

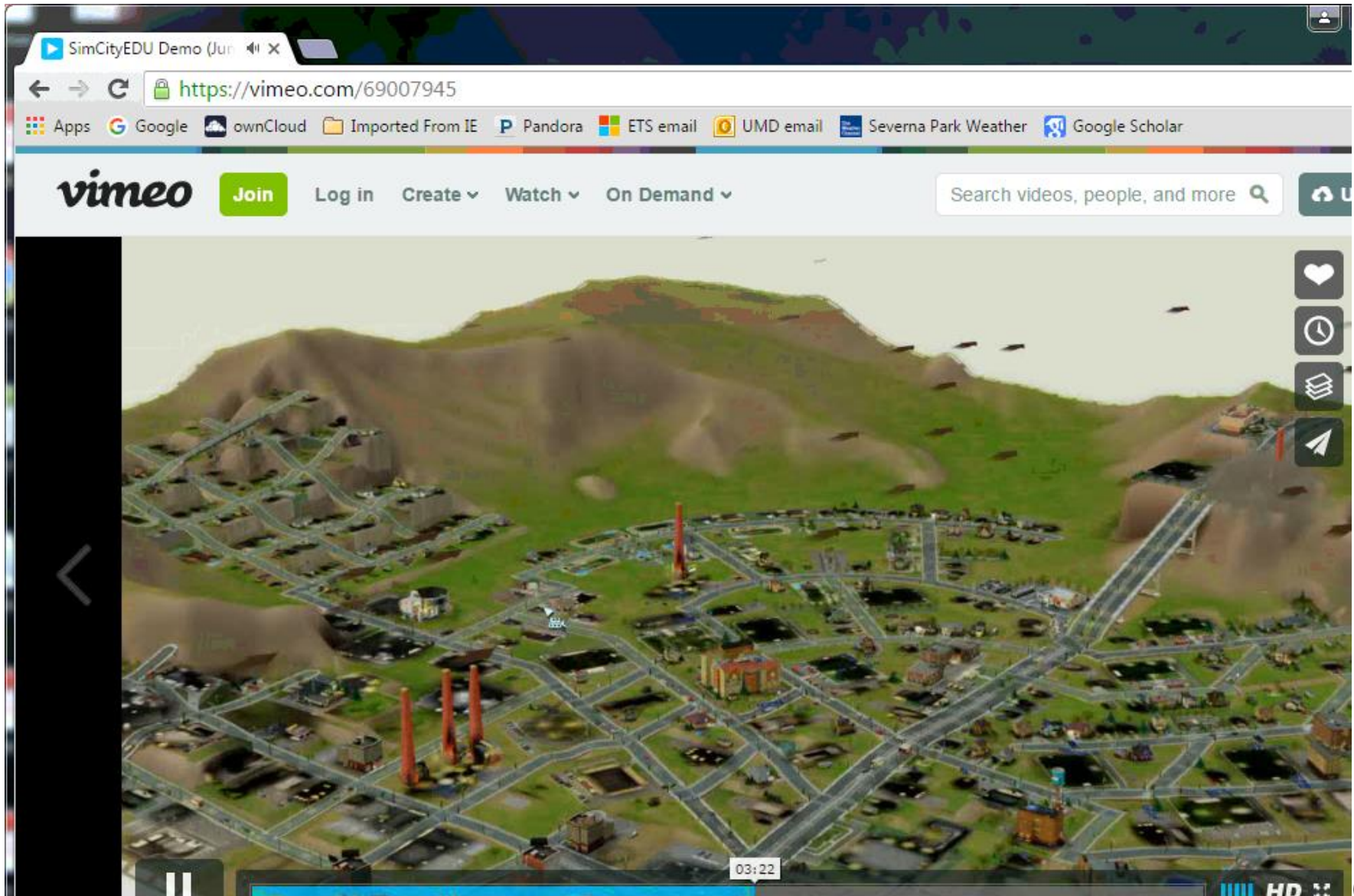
Messick (1994) on Assessment Design

- A construct-centered approach would begin by asking what complex of knowledge, skills, or other attributes should be assessed, presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society.
- Next, what behaviors or performances should reveal those constructs, and what tasks or situations should elicit those behaviors?
- Thus, the nature of the construct guides the selection or construction of relevant tasks as well as the rational development of construct-based scoring criteria and rubrics.

A snippet of SimCityEDU: Pollution Challenge!



A snippet of SimCityEDU: Pollution Challenge!

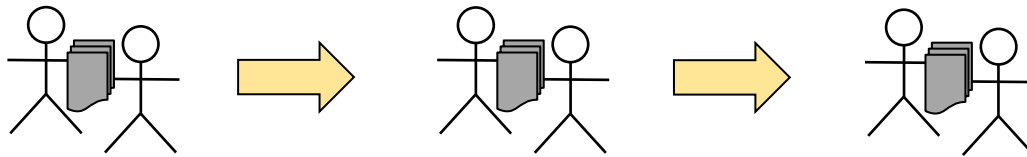


A snippet of SimCityEDU: Pollution Challenge!

```
<stateInfo>
  <Init Fragment="600" Style="blue" Language="ENG" />
  <FSMStates Count="7">
    <finishButtonFSM>Stop Flashing</finishButtonFSM>
    <popupSubmitButtonFSM>Stop Flashing</popupSubmitButtonFSM>
    <submitButtonFSM>Stop Flashing</submitButtonFSM>
    <continueButtonFSM>Stop Flashing</continueButtonFSM>
    <nextButtonFSM>Stop Flashing</nextButtonFSM>
    <timeoutFSM>Clear</timeoutFSM>
    <playmakerFSM>Finish</playmakerFSM>
  </FSMStates>
  <OtherVars Count="3">
    <F6_Response>A</F6_Response>
    <F6_Reason1>asdf</F6_Reason1>
    <F6_Reason2>asdf</F6_Reason2>
  </OtherVars>
</stateInfo>
<itemResult accessionNumber="TestAccNum" itemType="SBT" childItemAccessionNumber="176"
blockCode="TestBlockCode">
  <responseVariable cardinality="single" baseType="string">
    <candidateResponse>
      <value><![CDATA[{"Selection of relevant questions":"Y,Y,Y,N"}]]></value>
    </candidateResponse>
  </responseVariable>
  <responseVariable cardinality="single" baseType="string">
    <candidateResponse>
      <value><![CDATA[{"How far away is the well?":"(a)Yes","Follow-up":"(a)There is probably not
enough water underground"}]]></value>
    </candidateResponse>
  </responseVariable>
  <responseVariable cardinality="single" baseType="string">
    <candidateResponse>
      <value><![CDATA[{"Wells in other villages?":"(b)No","Follow-up":null}]]></value>
    </candidateResponse>
  </responseVariable>
  <responseVariable cardinality="single" baseType="string">
    <candidateResponse>
```



A Situative, Sociocognitive Perspective

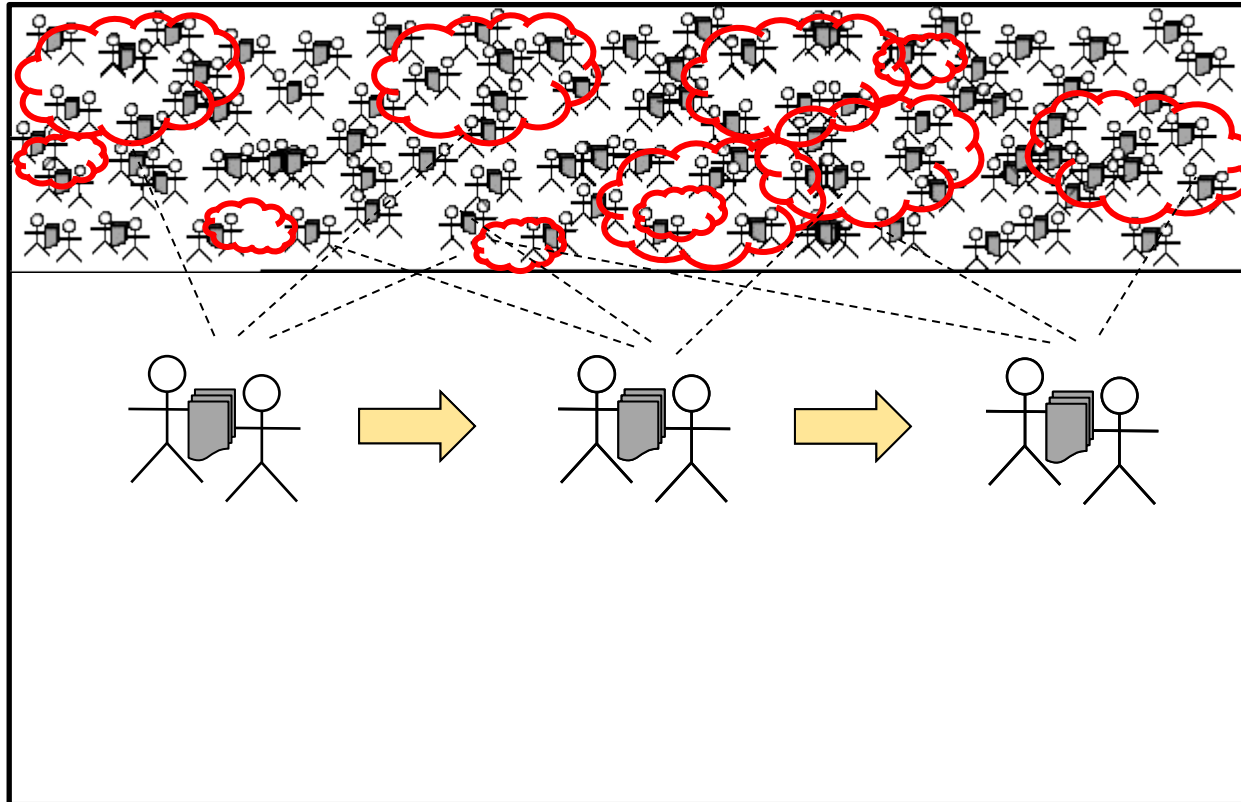


Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.



A Situative, Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

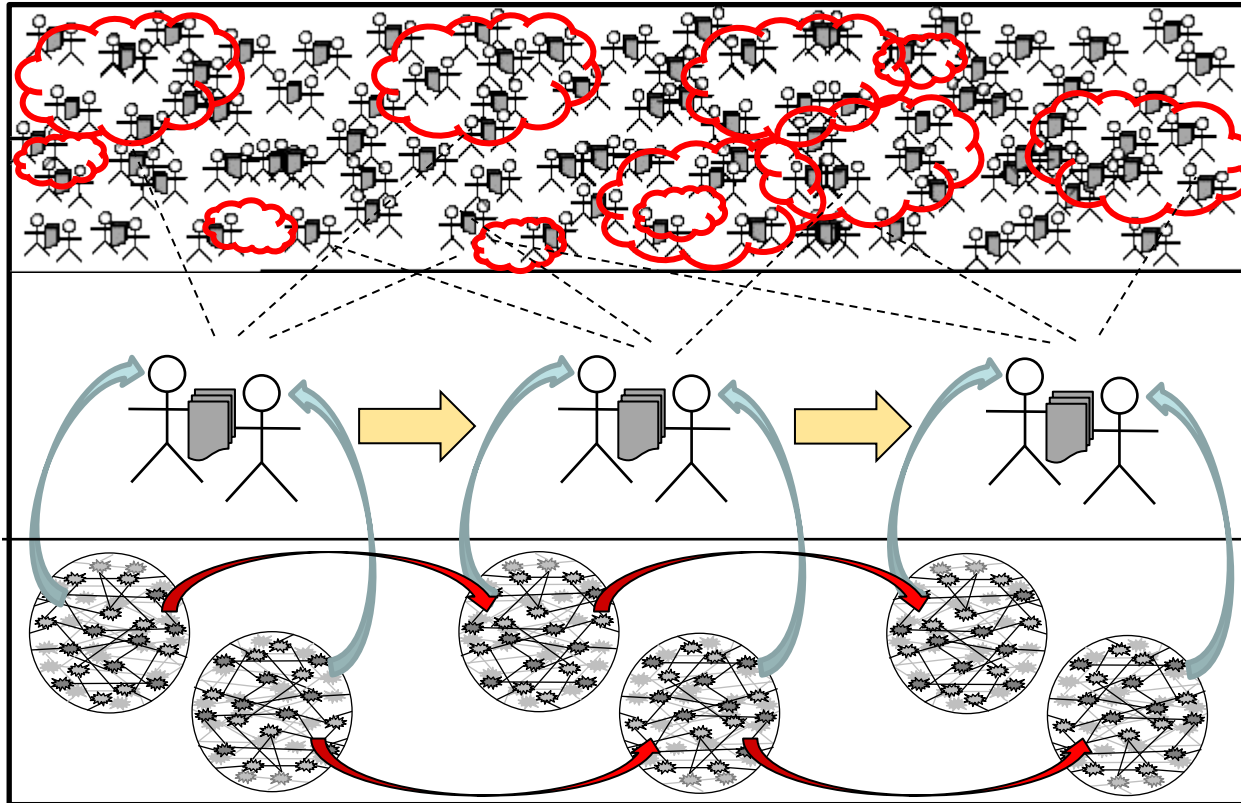
Regularities over unique instances of person-level situations and events.

Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.



A Situative, Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

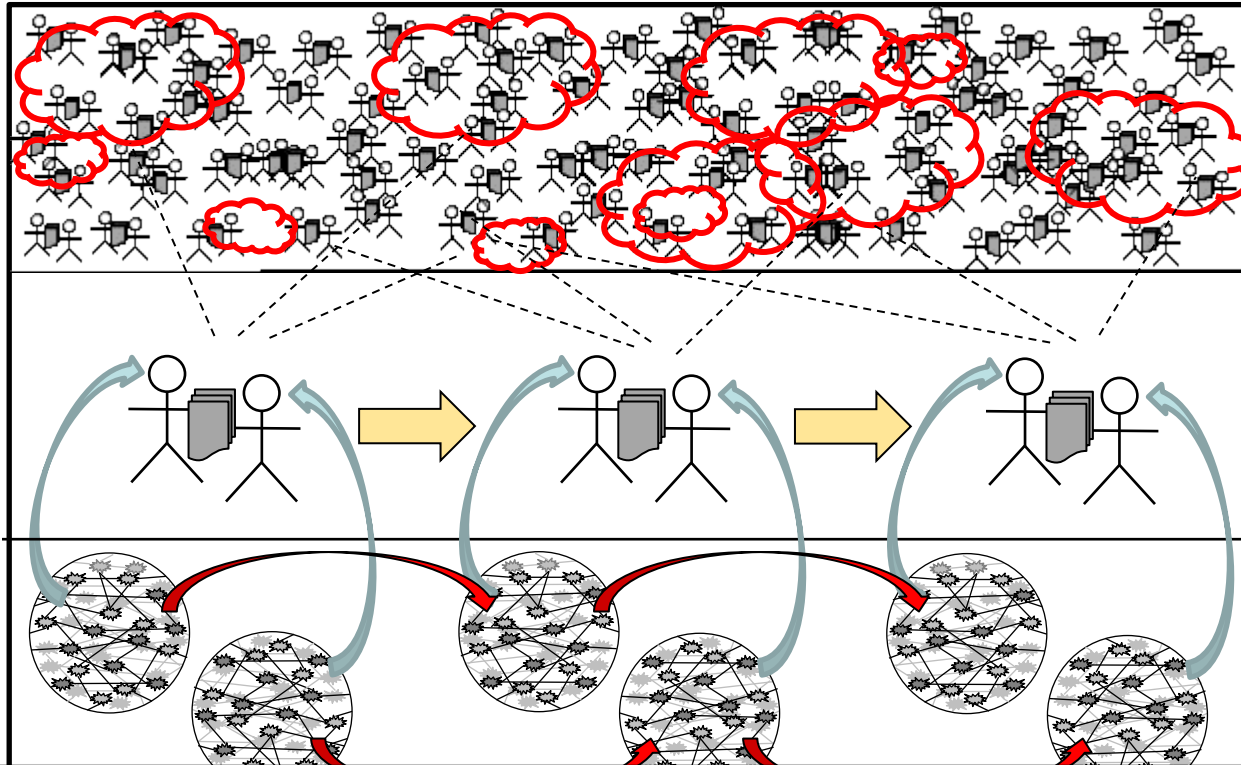
Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

A Situative, Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources**

developed through personal experiences.

Complex adaptive systems (CAS)

LCS patterns and cognitive resources are not 1-1

There are no measurement-model ϕ s.

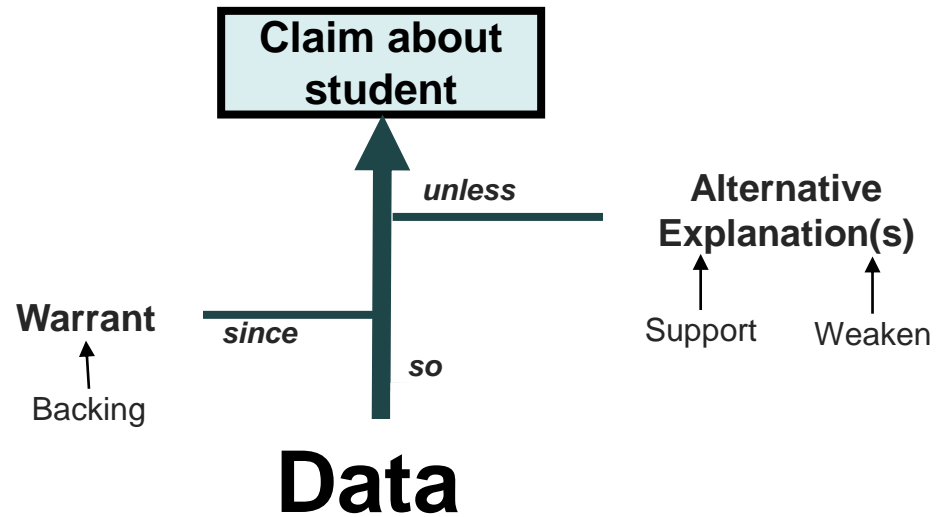


Some Implications for **Assessment**

- Every person-level situation builds around LCS patterns of many kinds and levels.
- An individual's experience of a situation assembles cognitive resources of many kinds, blended with features of that situation.
- The cognitive resources each person develops are unique. They depend on personal history, in a person's milieu of experience.
- Regularities *across* persons can arise due to similarities that shape the situations they have experienced. Thus arise patterns in people's resources and actions. There can be regularities and variation within and across people, within and across situations.
- Such regularities and variation in a set of situations (eg, tasks) as may arise – and as we may arrange to arise – are the grist of measurement modeling.

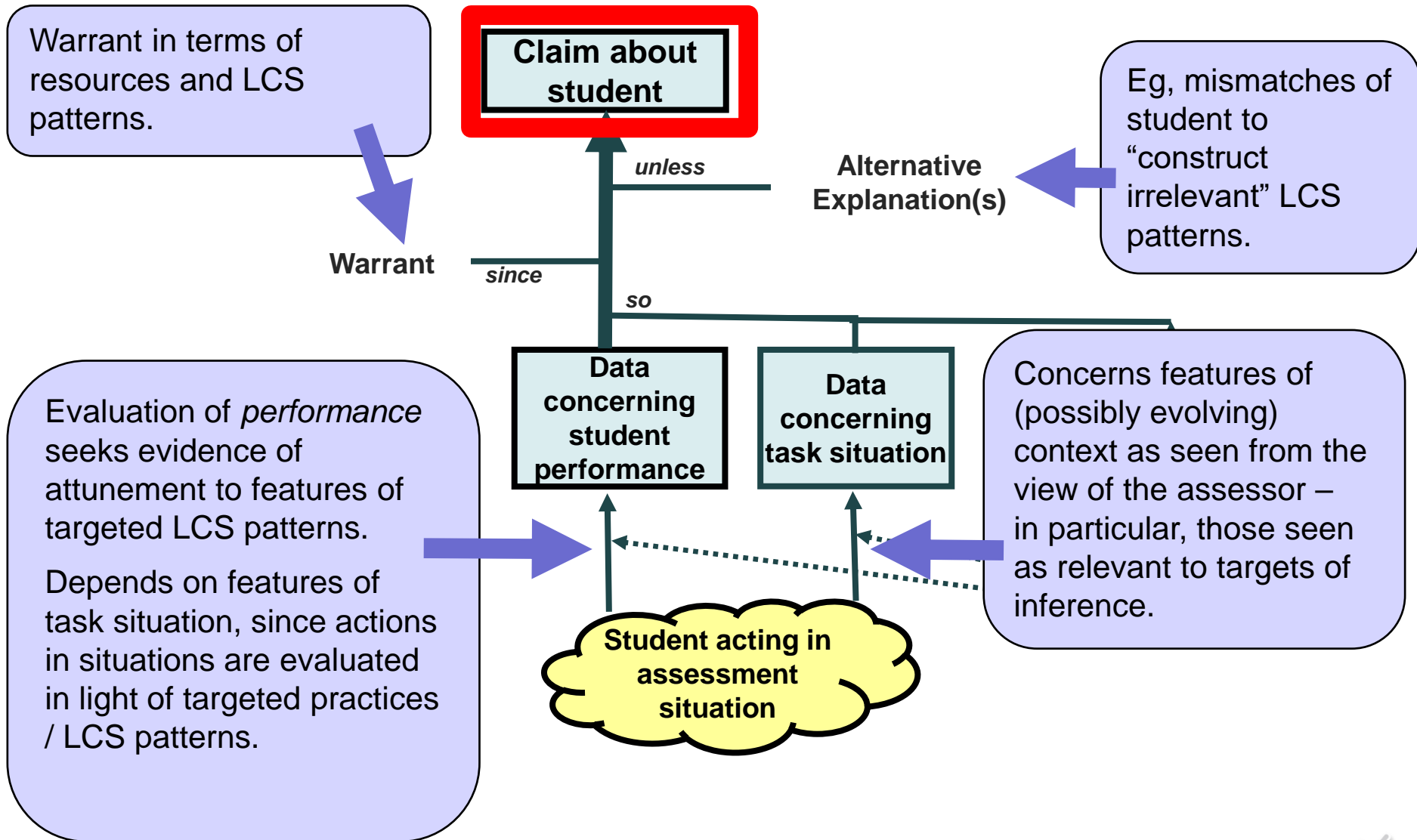
Evidentiary Argument in Educational Assessment

Evidentiary Argument in Educational Assessment

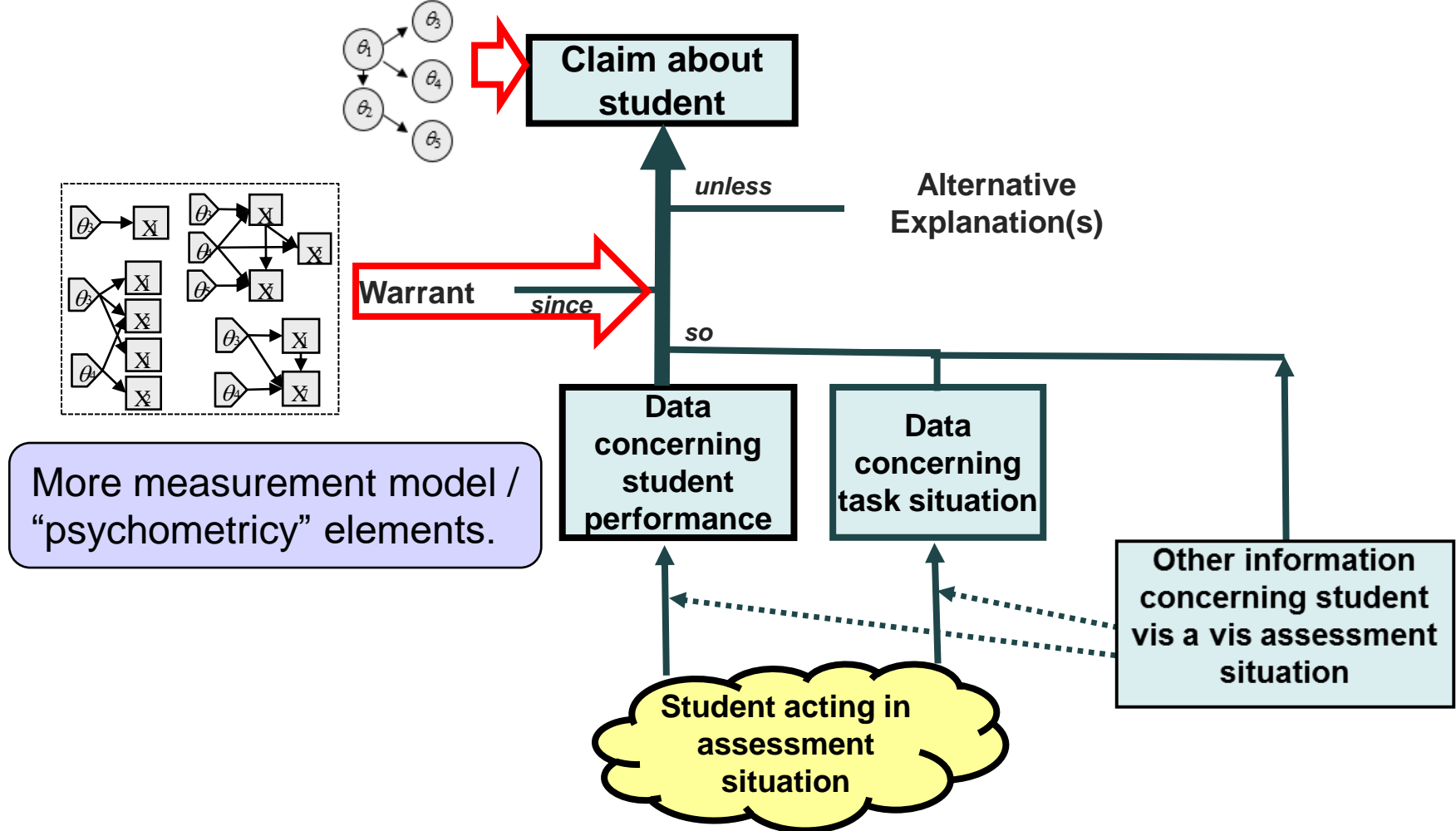


Warrant  *si*

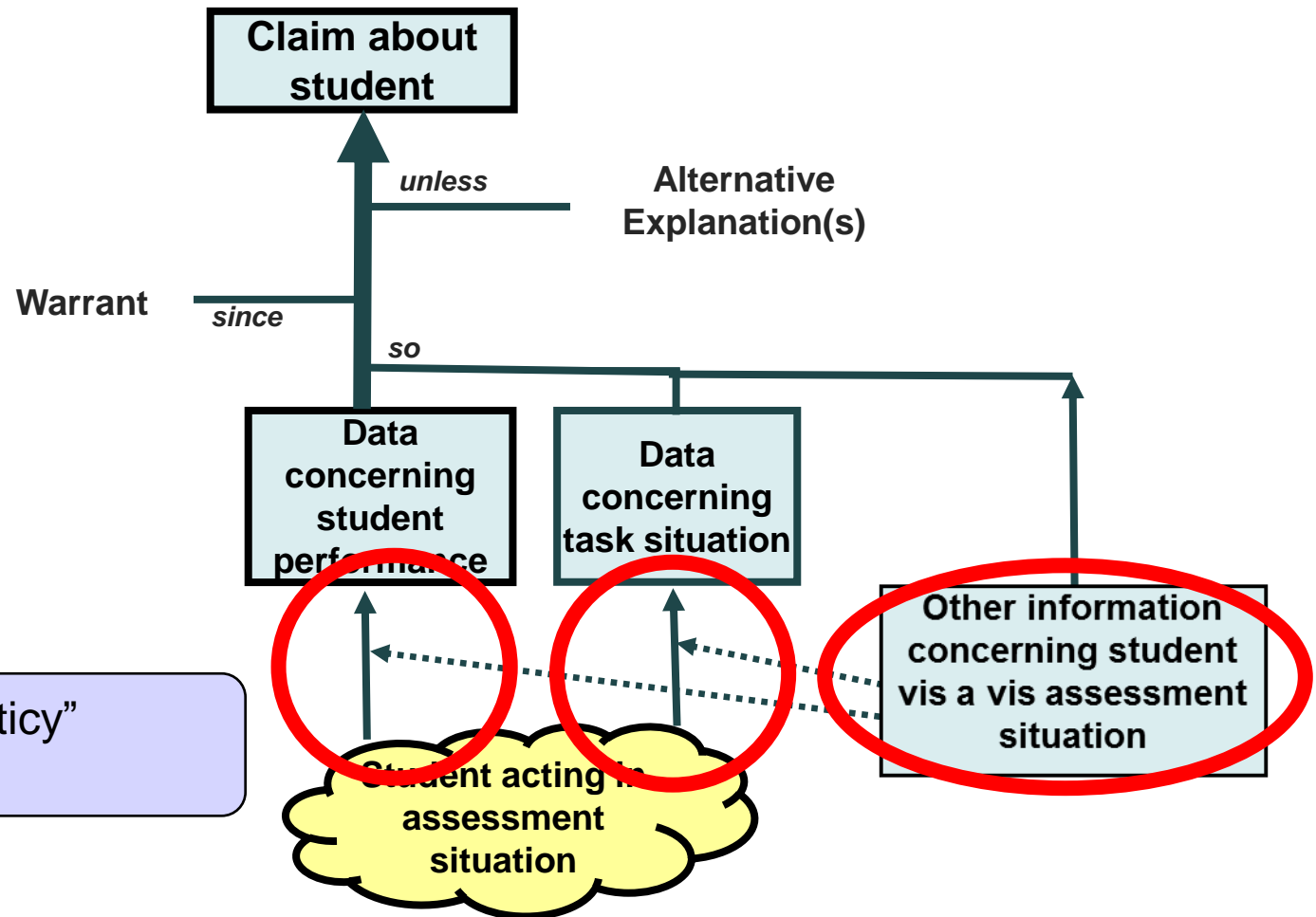
Evidentiary Argument in Educational Assessment



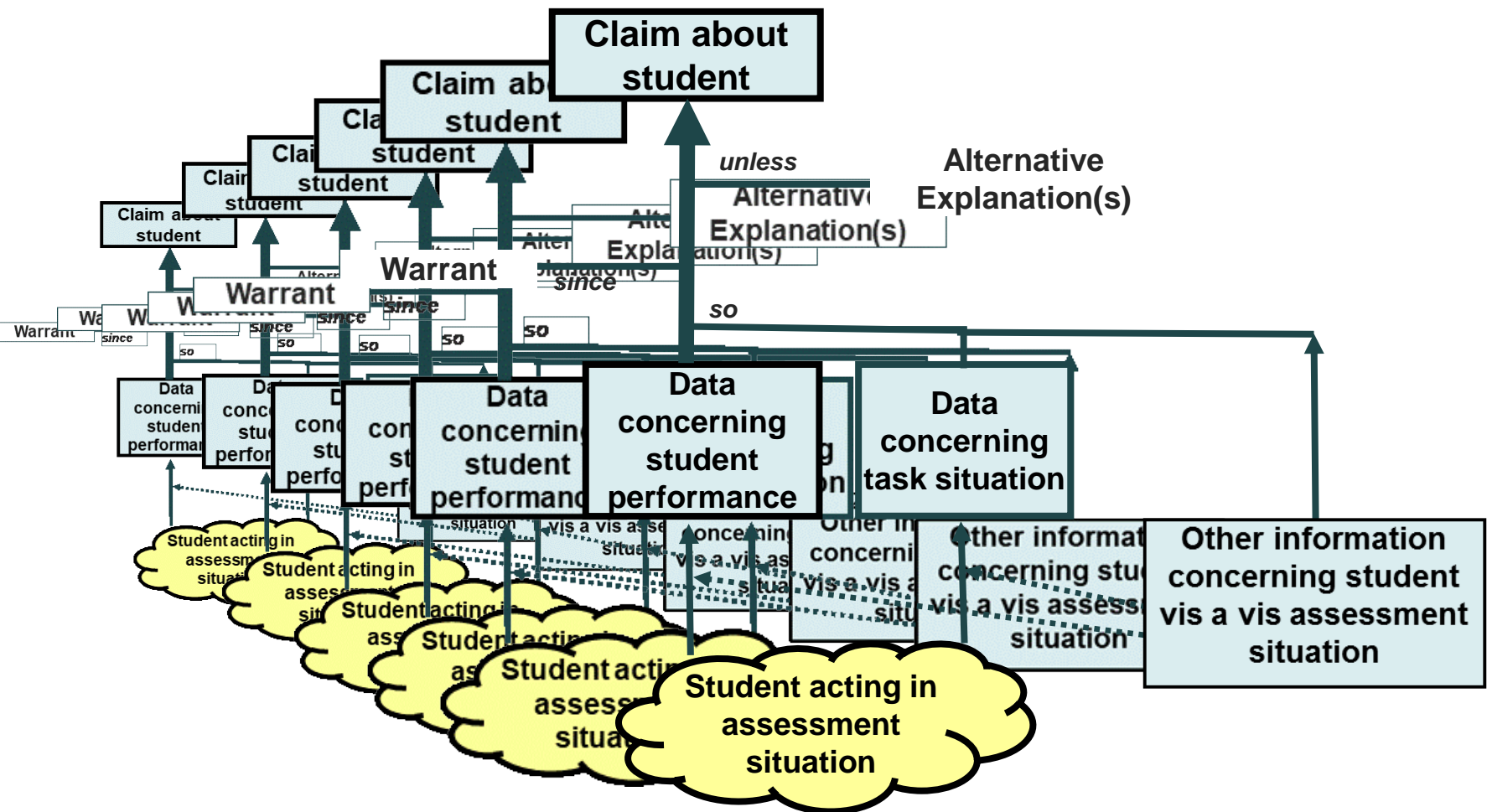
Evidentiary Argument in Educational Assessment



Evidentiary Argument in Educational Assessment



Evidentiary Argument in Educational Assessment

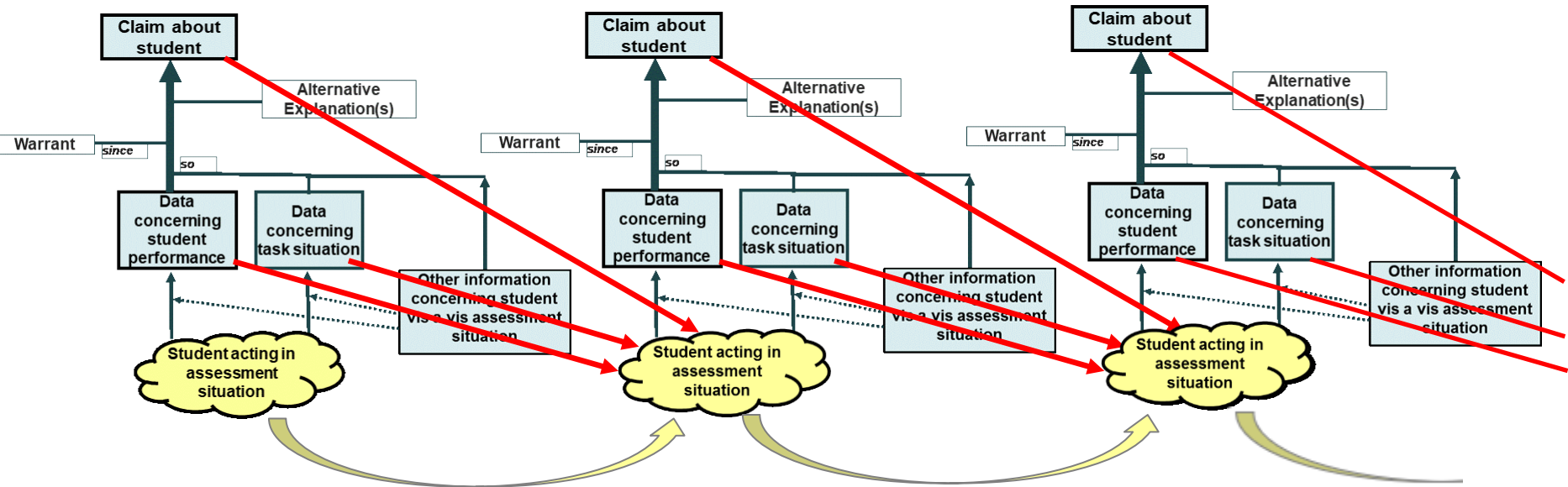


Extending the argument structure to interactive tasks such as simulations:

- Situation changes in response to student actions (and maybe other reasons)
- Interpretation must often account for certain past actions, situation features.



Evidentiary Argument in Educational Assessment

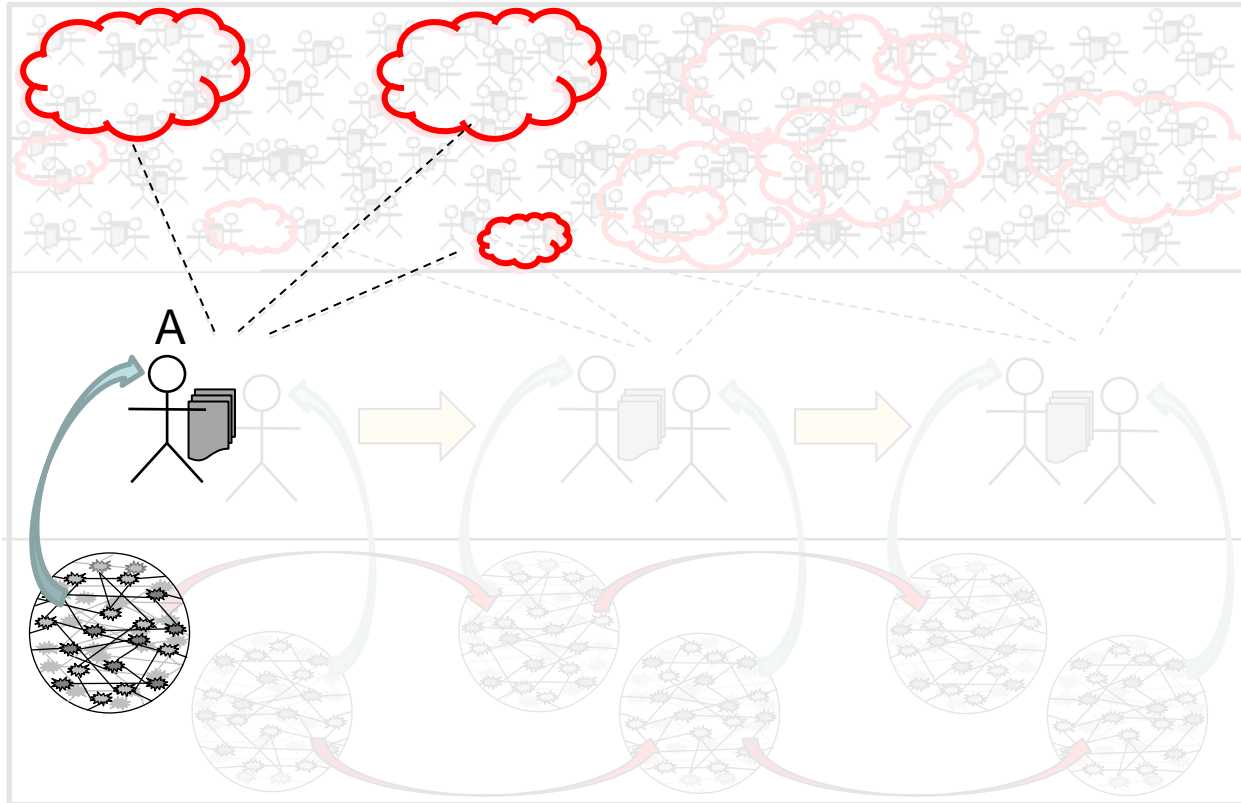


Extending the argument structure to interactive tasks such as simulations:

- Situation changes in response to student actions (and maybe other reasons)
- Interpretation must often account for certain past actions, situation features.



Measurement Models from a Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

Person-level experience.

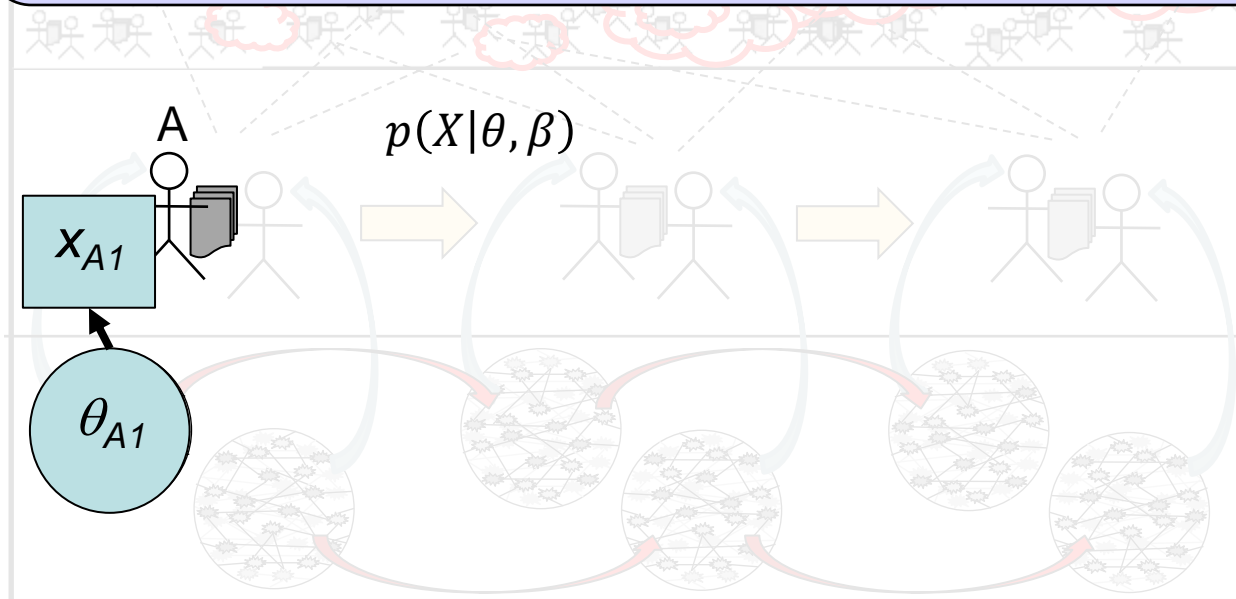
Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

Measurement Models from a Sociocognitive Perspective

Think of ed. measurement modeling as approximations within the overall patterns in some behavior domain within and across persons, that emerge in a given SC system.



situations and events.

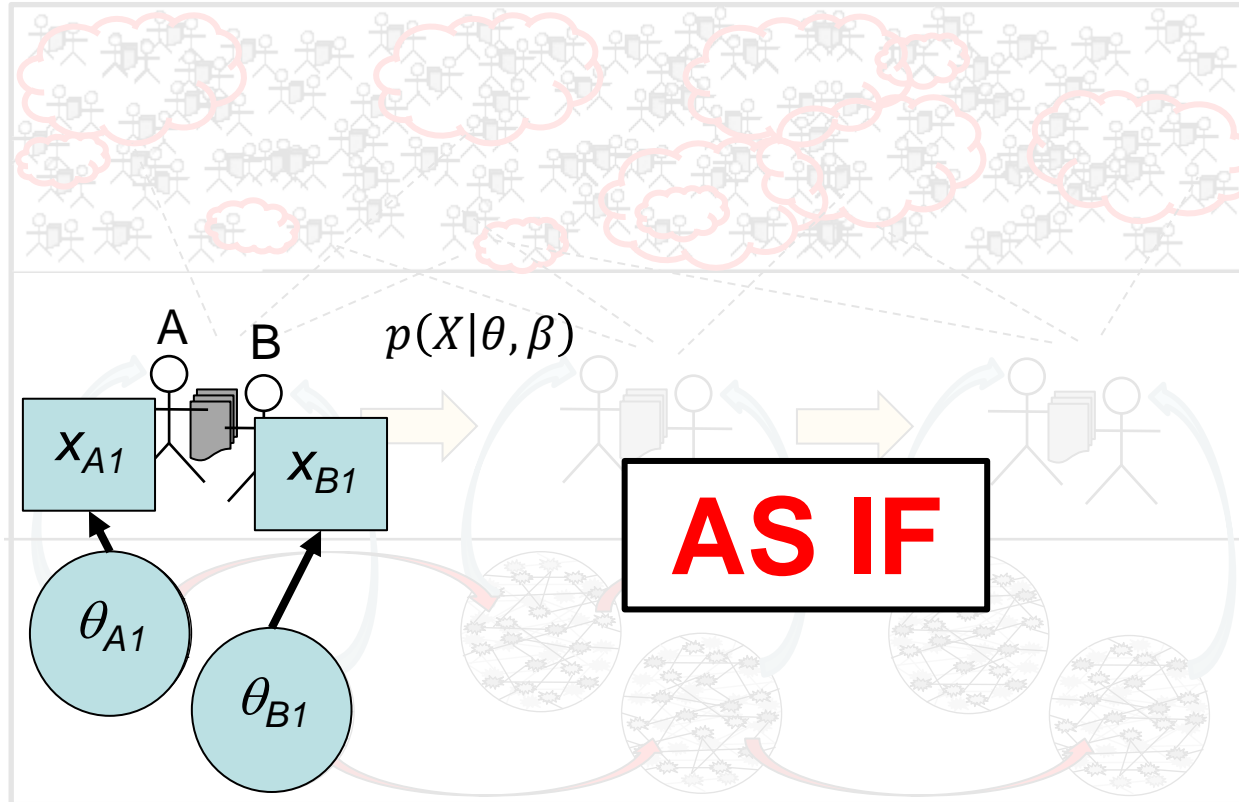
Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

Measurement Models from a Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

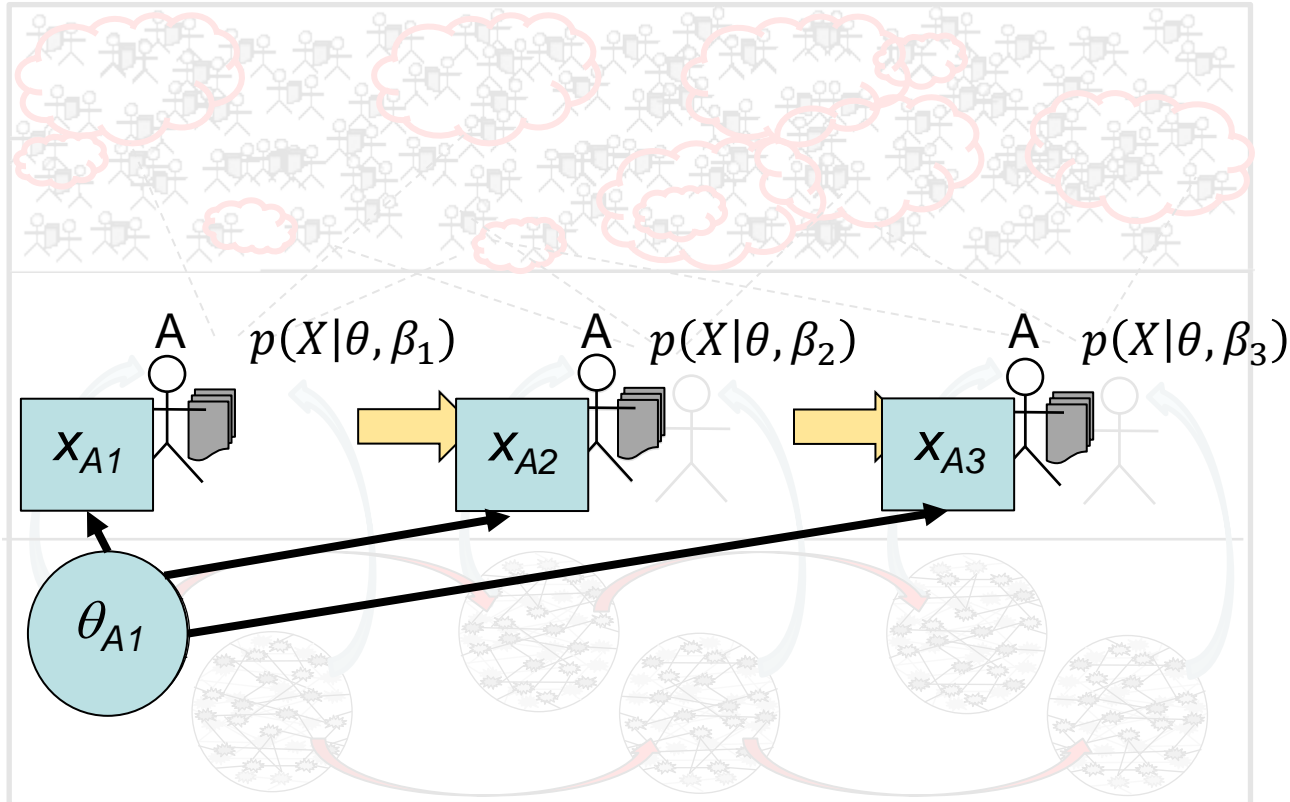
Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

Measurement Models from a Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

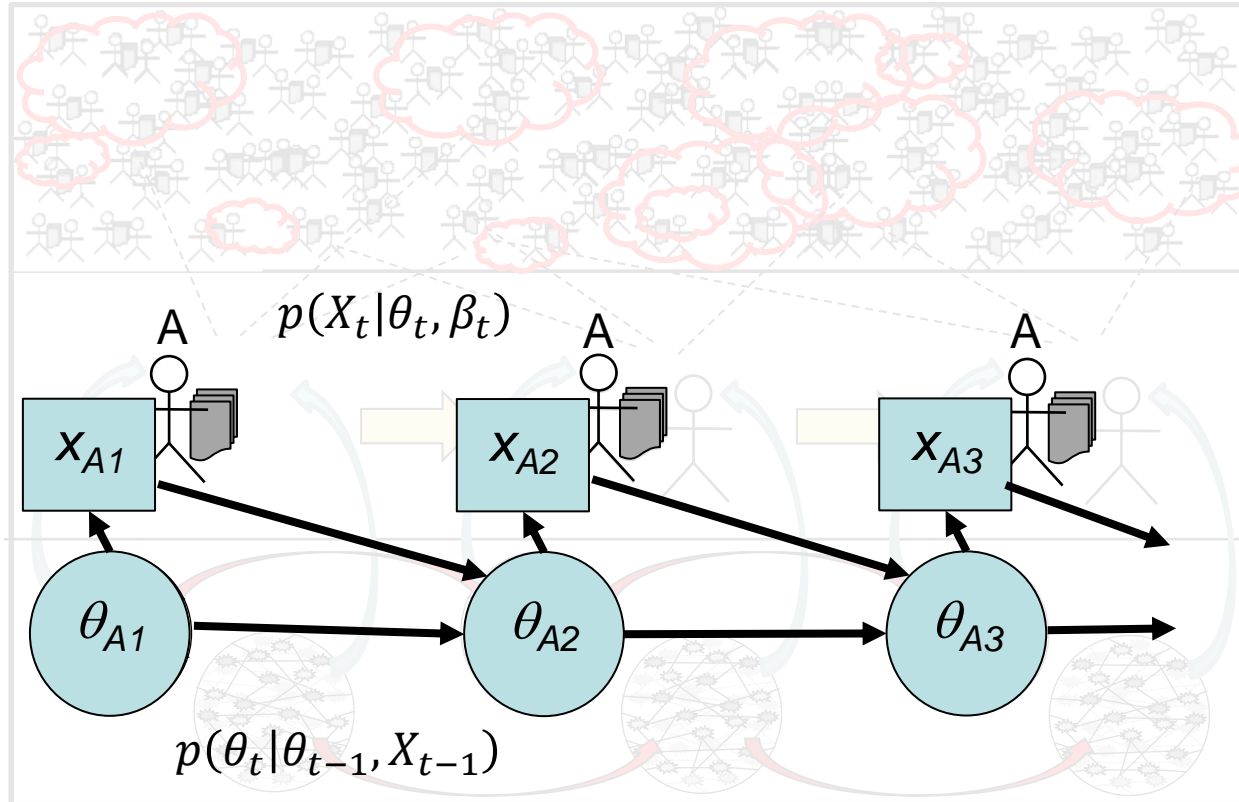
Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

Measurement Models from a Sociocognitive Perspective



Across-person linguistic, cultural, & substantive (LCS) patterns.

Regularities over unique instances of person-level situations and events.

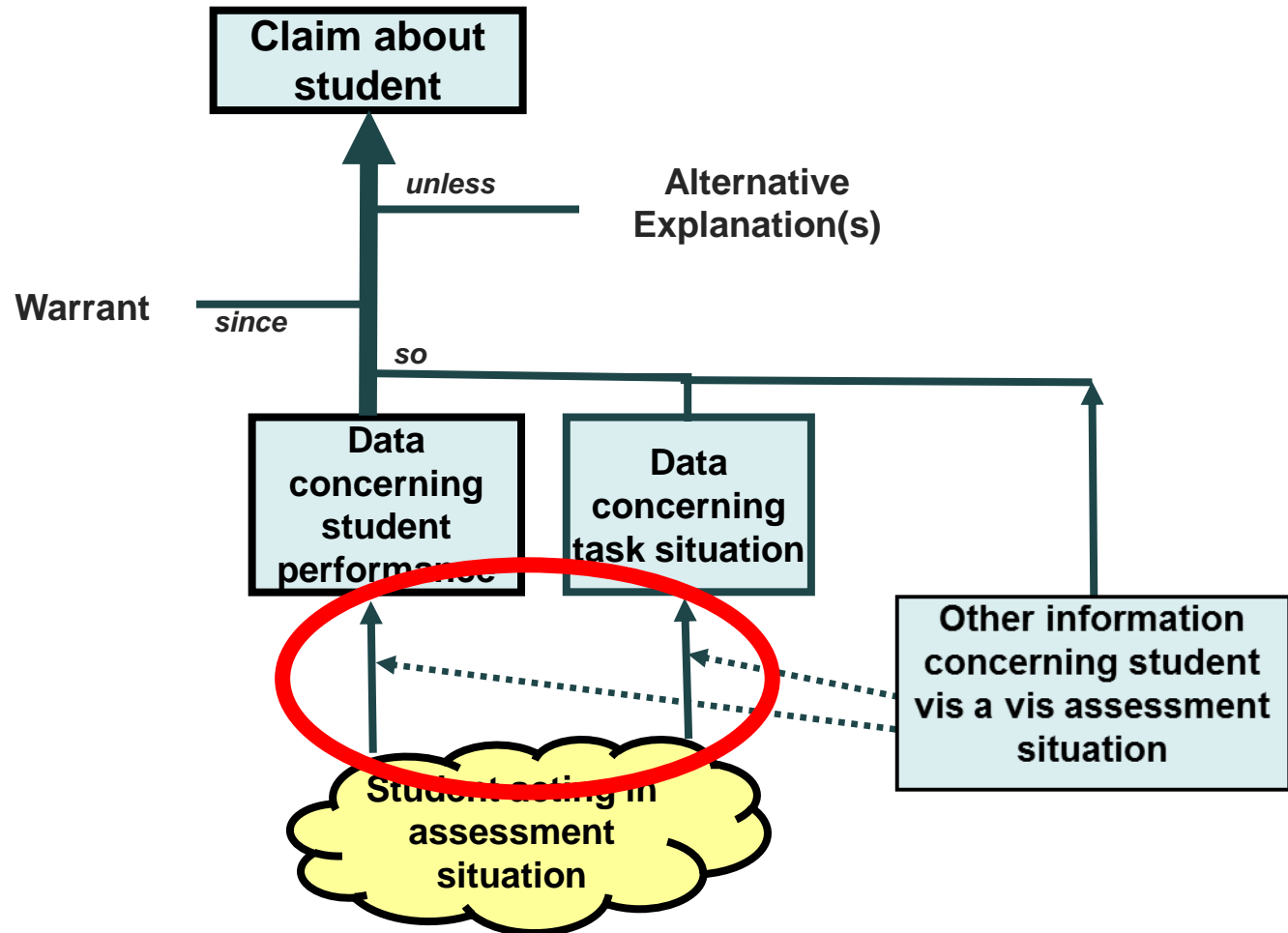
Person-level experience.

Situations and events as people experience them, interacting with the physical and social world.

Within-person sub-personal cognition.

Neural associations, unique patterns as **resources** developed through personal experiences.

Evidence Identification in Educational Assessment

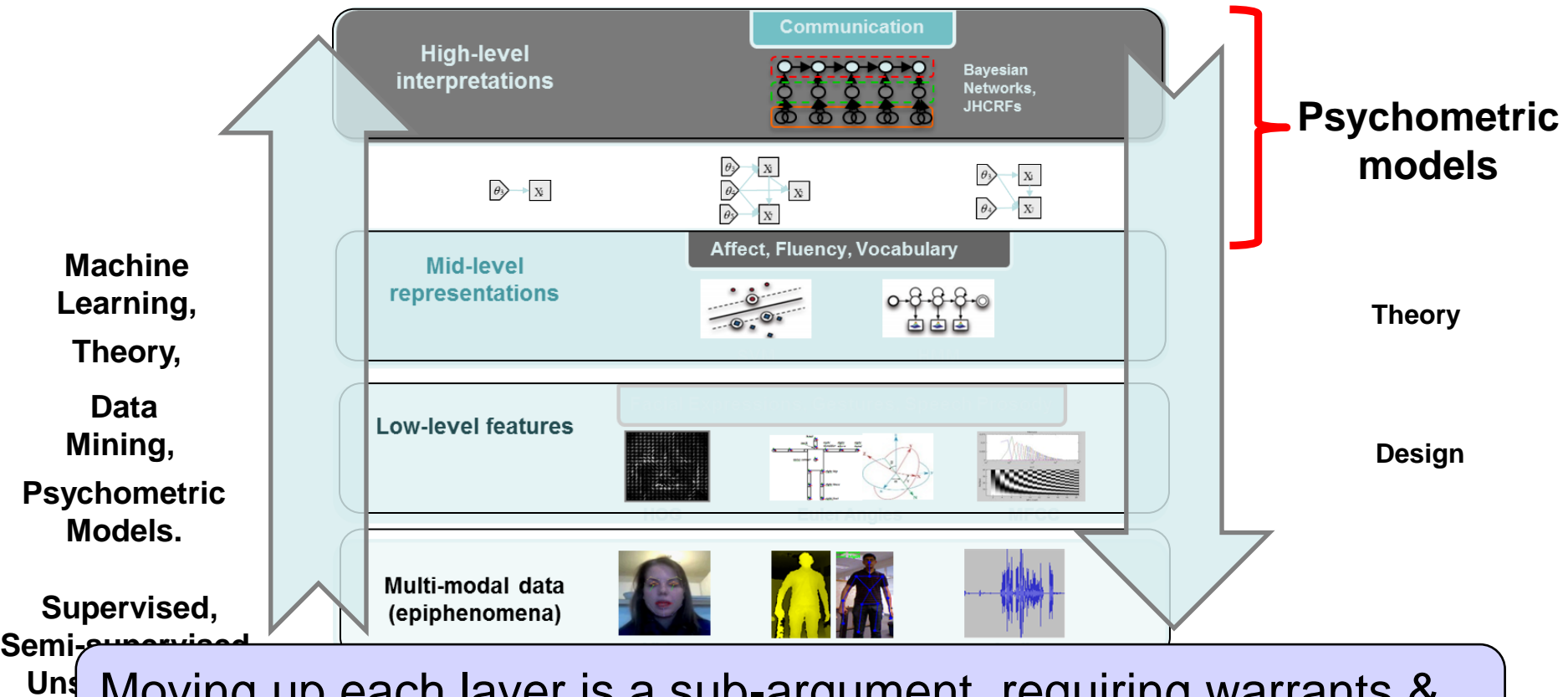


SimCityEDU: Pollution Challenge!

Feature Evaluation

Evidence for constructs from low-level data.

Hierarchies of chain of evidentiary reasoning (can be up & down, theory-aided.)



Moving up each layer is a sub-argument, requiring warrants & backing, and subject to alternative explanations to examine.

SimCityEDU: Pollution Challenge!

Feature Evaluation

Evidence for constructs from
Hierarchies of chain of evidence

Construct was levels on a systems-thinking learning progression variable – reflects *kinds* of things people can do in *kinds* of situations. Backbone across challenges.

High-level interpretations



JHCRFs

Psychometric models



Mid-level representations

Summary functions of these patterns were input (X's) into a Bayes net measurement model.

Strategic patterns of verb clauses; e.g., build low-pollution plant before bulldozing high-pollution one.

Low-level features

“Verb clauses”: Locations, times, durations and objects of meaningful actions like “rezone,” “bulldoze,” “query map.”

Multi-modal data (epiphenomena)

Interface and situation-specific locations, times, and durations of clicks, hovers, drag & drops, etc.

Machine Learning,
Theory,

Data Mining,
Psychometric Models.

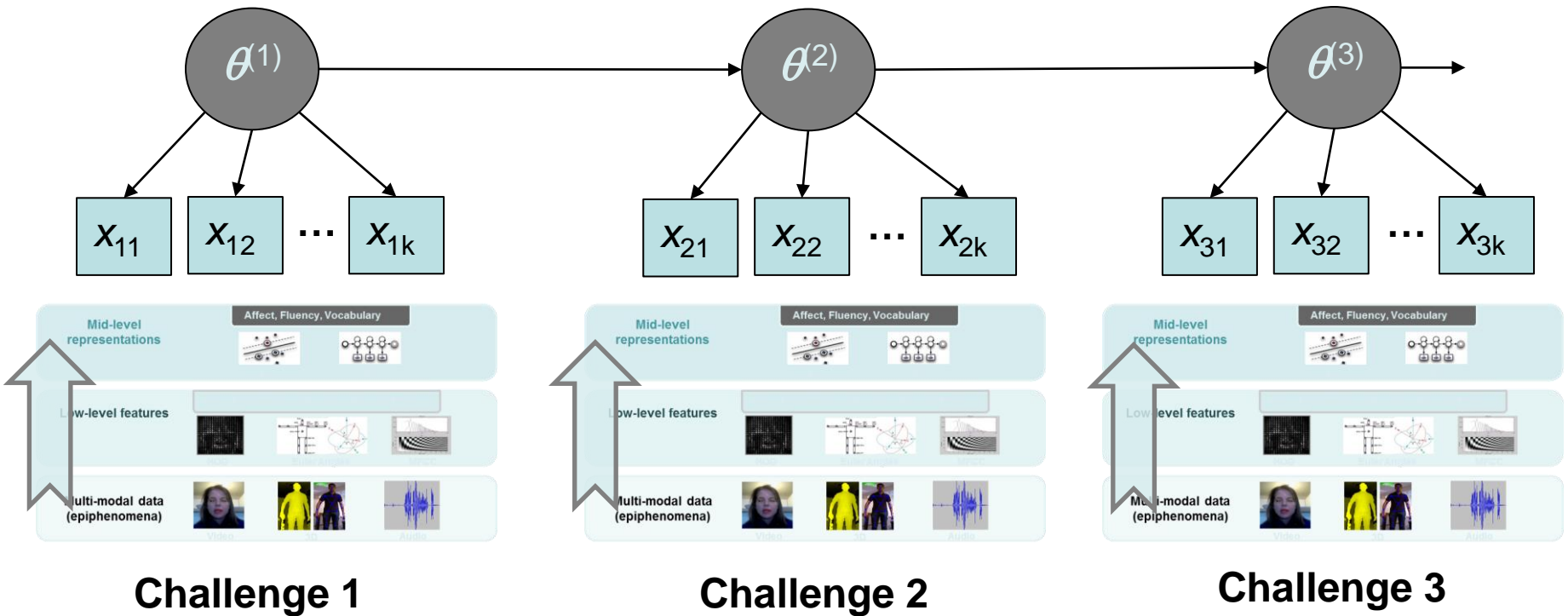
Supervised,
Semi-supervised,
Unsupervised

from von Davier, Khan, & Kerr

SimCityEDU: Pollution Challenge!

High-level
interpretations

Levels on systems-thinking
learning progression



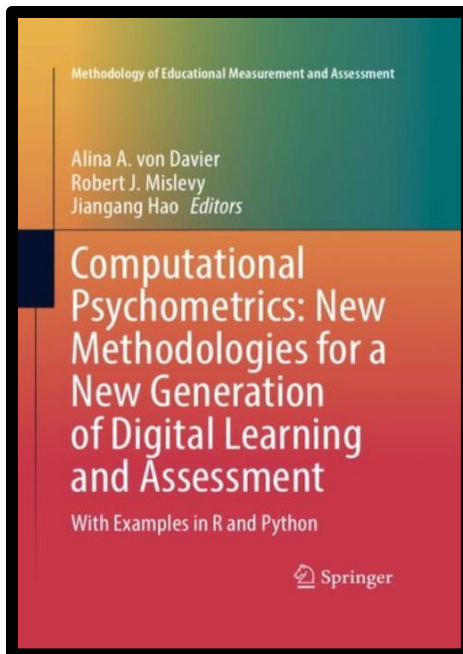
Advantages of Measurement Modeling at Higher Layers

- Synthesize evidence across tasks/evolving situations in terms of evidence about constructs. Not really about “scores.”
- Validity: Methods for examining quality of “as if” reasoning about constructs (history from G-theory, multi-trait multi-method).
- Reliability: Precision of estimation provides metric for tuning / critiquing / comparing evidence identification methods.
- Fairness: Methods for analyzing interaction of analytic methods with variables such as language & culture as to impact on inferences. Guide revision of tasks, methods, or extended models to deal with them when needed.

Concluding Remarks

- Psychological/social underpinning and substance of an assessment are essential to interpreting and using measurement-model elements (even if implicit).
- A situative, sociocognitive / complex systems perspective connects disciplines involved in learning and assessment. It subsumes trait, behavioral, and cognitive perspectives.
- Measurement modeling remains useful in designing, critiquing, and using assessments, for managing issues of evidence and inference. But it is not sufficient.
- Argumentation structuring provides a framework for working through these issues. It subsumes the measurement modeling frame.

Some References



Behrens, J.T., Mislevy, R.J., DiCerbo, K.E., & Levy, R. (2012). An evidence centered design for learning and assessment in the digital world. In M.C. Mayrath, J. Clarke-Midura, & D. Robinson (Eds.), *Technology-based assessments for 21st century skills* (pp. 13-54). Charlotte, NC: Information Age.

Byrne, D. (2002). *Interpreting quantitative data*. London: Sage Publications.

Ellis, N. C., & Larsen-Freeman, D. (eds.) (2009). *Language as a complex adaptive system*. Oxford: Wiley.

Mislevy, R.J. (2019). Advances in the science of measurement and cognition. *Annals of the American Academy of Politics & Social Science*, 683(1), 164-182.

Hammer, D., Elby, A., Scherr, R.E., & Redish, E.F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89-120). Greenwich, CT: Information Age Publishing.

Mislevy, R.J. (2016). How developments in psychology and technology challenge validity argumentation. *Journal of Educational Measurement*, 53, 265-292.

Mislevy, R.J. (2017). Resolving the paradox of rich performance tasks. In H. Jiao & Lissitz, R. W. (Eds.), *Test fairness in the new generation of large-scale assessment* (pp. 1-46). Charlotte, NC: Information Age Publishing.

Mislevy, R.J. (2018) *Sociocognitive foundations of educational measurement*. NY: Routledge.

Mislevy, R.J. (2018). On integrating psychometrics and learning analytics in complex assessments. In H. Jiao & Lissitz, R. W. (Eds.), *Test data analytics and psychometrics: Informing assessment practices* (pp. 1-48). Charlotte, NC: Information Age Publishing.

Mislevy, R.J., Yan, D., Gobert, J., & Sao Pedro, M. (2020). Automated scoring with intelligent tutoring systems. In D. Yan, A.A. Rupp, & P.W. Foltz (Eds.), *Handbook of automated scoring: Theory into practice* (pp. 403-422). Boca Raton, FL: CRC Press/Routledge.

Mislevy, R.J. (2020). An evidentiary-reasoning perspective on automated scoring: Commentary on Chapters 1-6. In D. Yan, A.A. Rupp, & P.W. Foltz (Eds.), *Handbook of automated scoring: Theory into practice* (pp. 151-169). Boca Raton, FL: CRC Press/Routledge.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell.

Von Davier, A., Mislevy, R.J., & Hao, J. (in press). *Computational psychometrics: New methodologies for a new generation of digital learning and assessment*.

Yan, D., Rupp, A.A., & Foltz, P.W. (Eds.), *Handbook of automated scoring: Theory into practice*. Boca Raton, FL: CRC Press/Routledge.

Thank you!