

The Challenge of Context

Robert J. Mislevy

Educational Testing Service
National Center for Research on Evaluation, Standards, and Student Testing

Assessment for Learning vs. Assessment for Accountability

- An overly stark dichotomy
- Different configurations to meet different purposes, under different constraints.
- The role of *context* is key to understanding.

The California Learning Assessment System (CLAS)

- On-demand assessment
- Embedded assessment
- Organic portfolios

AP Studio Art Portfolio Assessment

- Work created throughout the year, usually part of a class
- Standards-based
- Sparse common language, rich local languages

Assessment as Evidentiary Reasoning

- What we really want to know about is...
 - what students know, can do, should work on next.
- What we get to work with is...
 - We see them say, do, or make something in a handful of particular situations.

Principles from Evidentiary Reasoning, I

Evidence forming the basis for probabilistic conclusions has three major properties that must be established:

- relevance
- credibility
- inferential weight.

Kadane & Schum, 1996

Principles from Evidentiary Reasoning, II

The evidentiary value of data depends on...

- The question you are asking
- what else you know to condition your interpretation of the data

A little more about “The question you are asking”

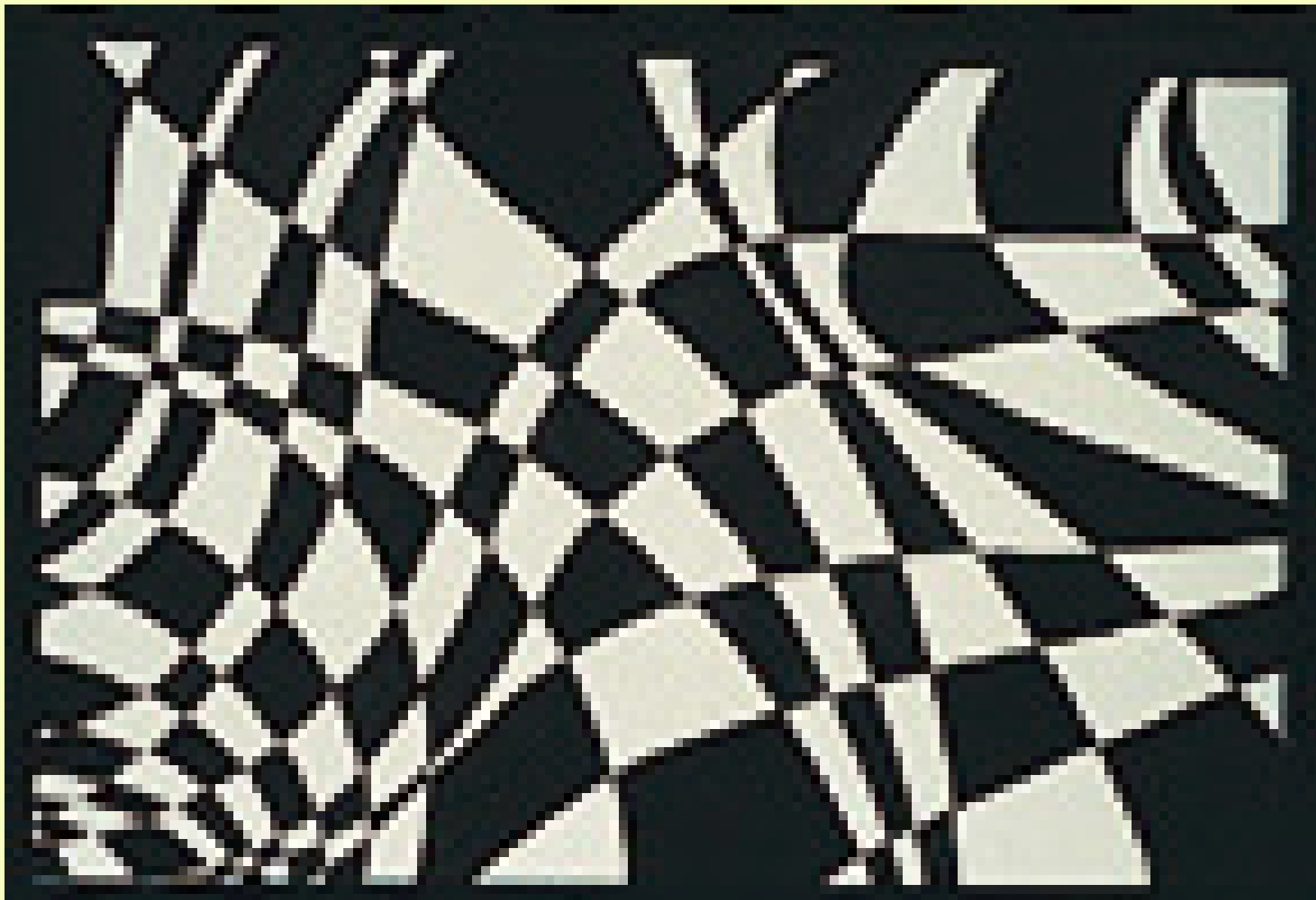
- “Frames of discernment”
- Relevance to Assessment

A little more about “What else you know”

- The ACTFL Reading Proficiency Scale
- AP Studio Art Concentration section









A little more about “What else you know”

- The ACTFL Reading Proficiency Scale
- AP Studio Art Concentration section
- CLAS On-Demand vs. Embedded

A further consideration: What you can ask *of* students

- Expectations / standards of evaluation
- Mechanics of assessment /
Representation forms

The Challenge

- Can we retain the relevance and connectedness that is associated with classroom assessment,
- and at the same time provide the communication and credibility for distant users that is associated with low-context on-demand tests?

Conclusion

- Multiple components, with different evidentiary characteristics, to support multiple purposes.
 - Different relationships between data and the amount of context that has to be agreed on and shared.
 - Different degrees of contextualized /conditional interpretation of performance.