

The Fundamental Processes in Assessment Delivery



Linda Steinberg Russell Almond Robert Mislevy

Educational Testing Service

prepared for the QTI Symposium, Tokyo, July, 2001

Of-heard words and phrases



- Computer-based / Web-based
- Standardized test
- Multiple-choice / Performance Assessment
- High-stakes / Accountability
- Standards-based
- Norm-referenced vs. Criterion referenced
- Adaptive test

The everyday language of testing

- Evolved for particular assessment configurations
- Focus on surface features
- Confounds purposes for assessment with methods for delivery & analysis
- Inadequate for describing important distinctions or designing assessments

Schemas for assessment

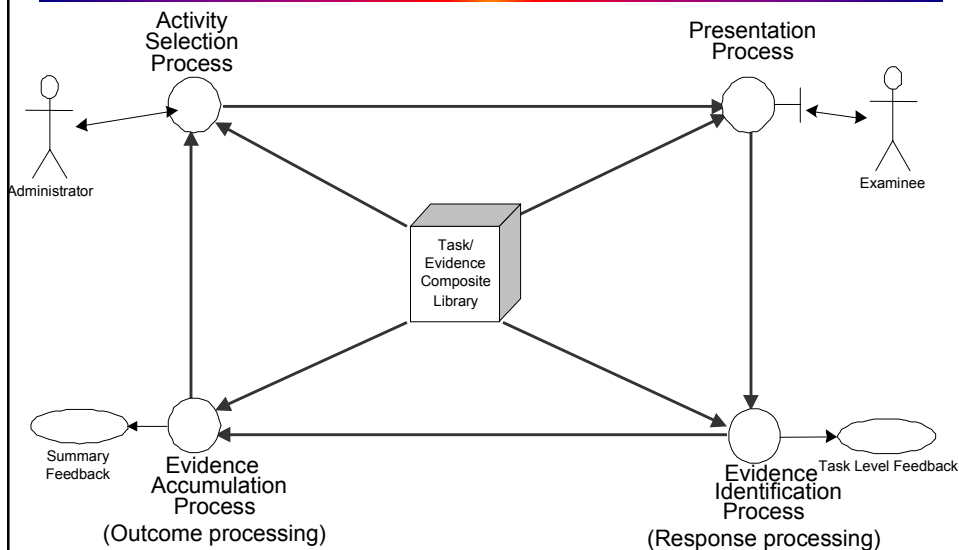
- Social context
 - Purpose
 - Constraints
 - Relationship to instruction / job / etc.
- Evidence-Centered Assessment Design
 - Explicating the evidentiary argument
 - Messick quote next
- *Four-process delivery system (ETS/IMS)*

Evidence-Centered Design

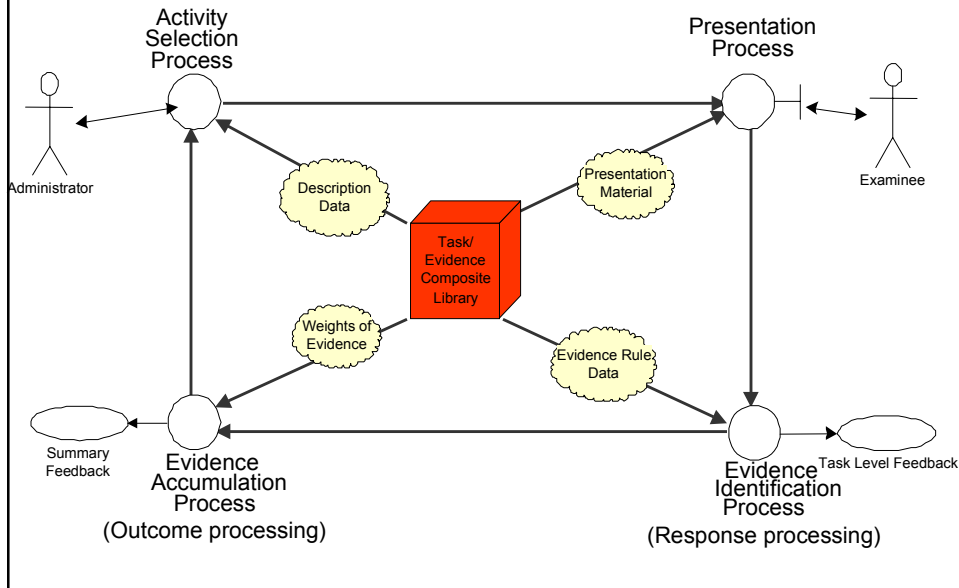
- [W]hat complex of knowledge, skills, or other attribute should be assessed, presumably because they are tied to ... objectives of instruction or are otherwise valued.
- Next, what behaviors or performances should reveal those constructs, and
- what tasks or situations should elicit those behaviors?

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.

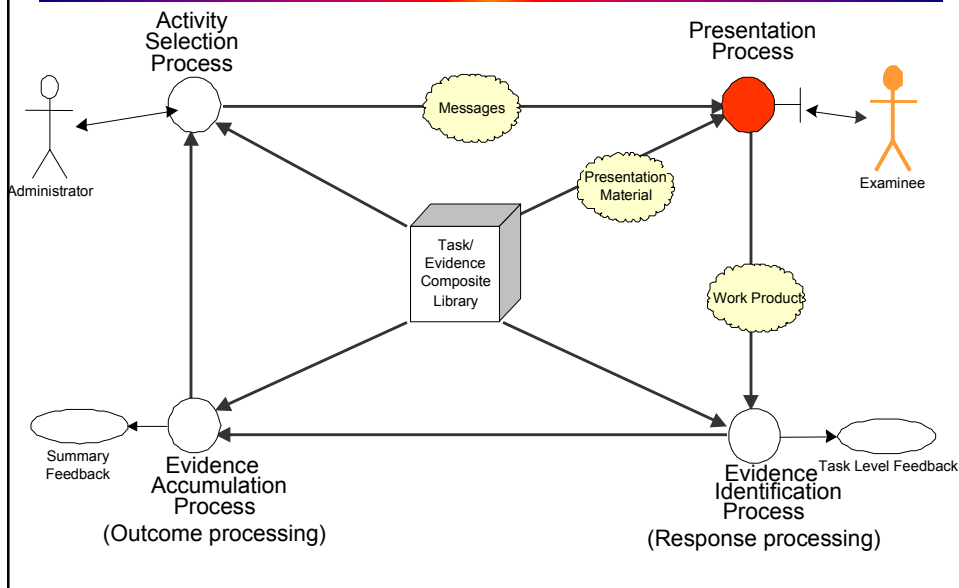
A Four-Process Delivery System



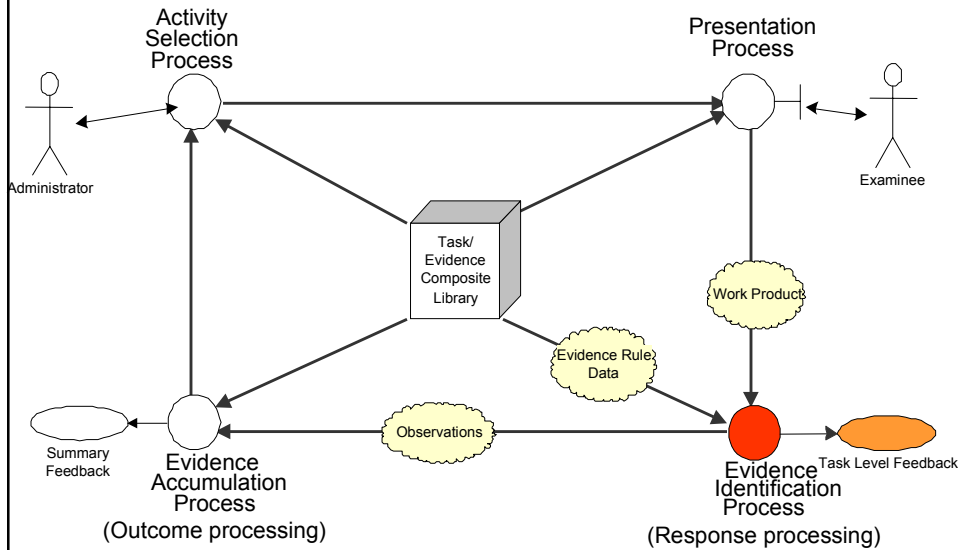
A Four-Process Delivery System



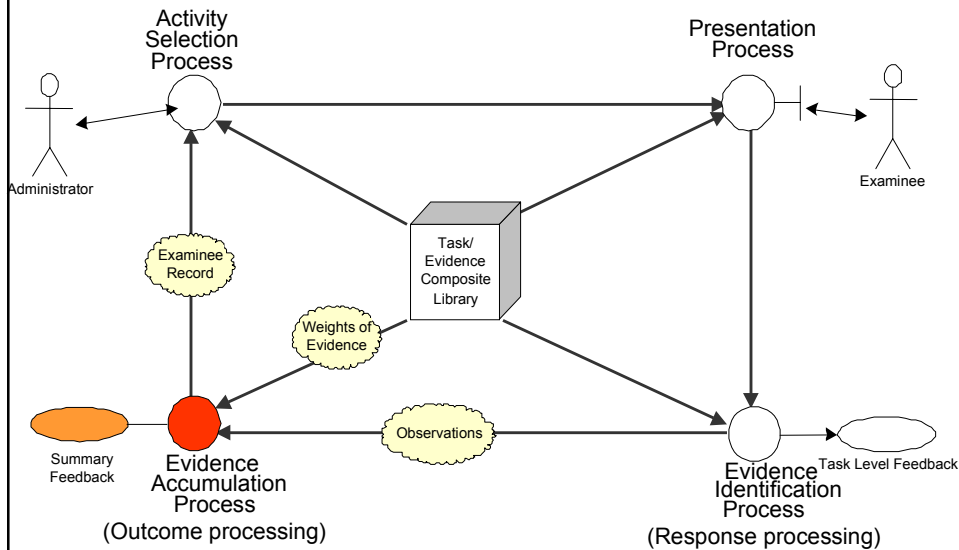
A Four-Process Delivery System



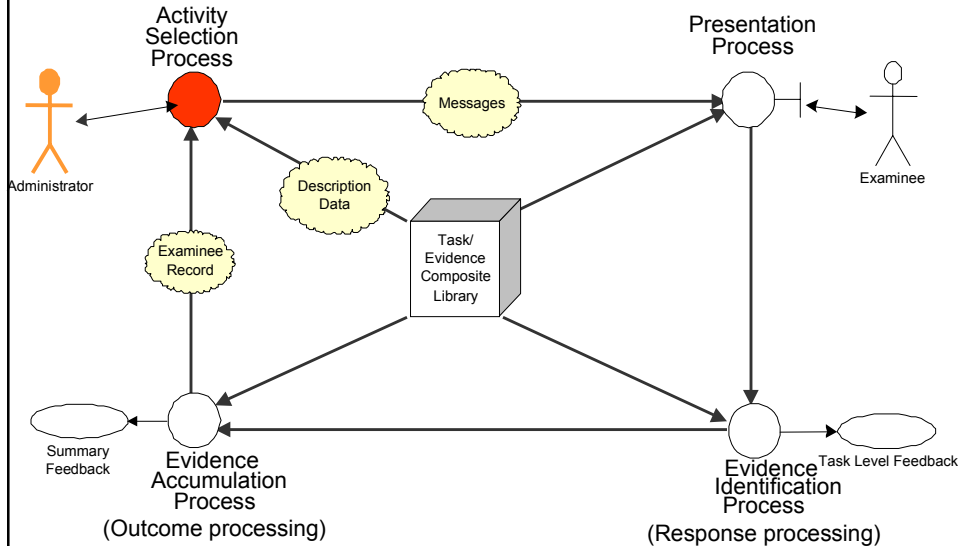
A Four-Process Delivery System



A Four-Process Delivery System



A Four-Process Delivery System

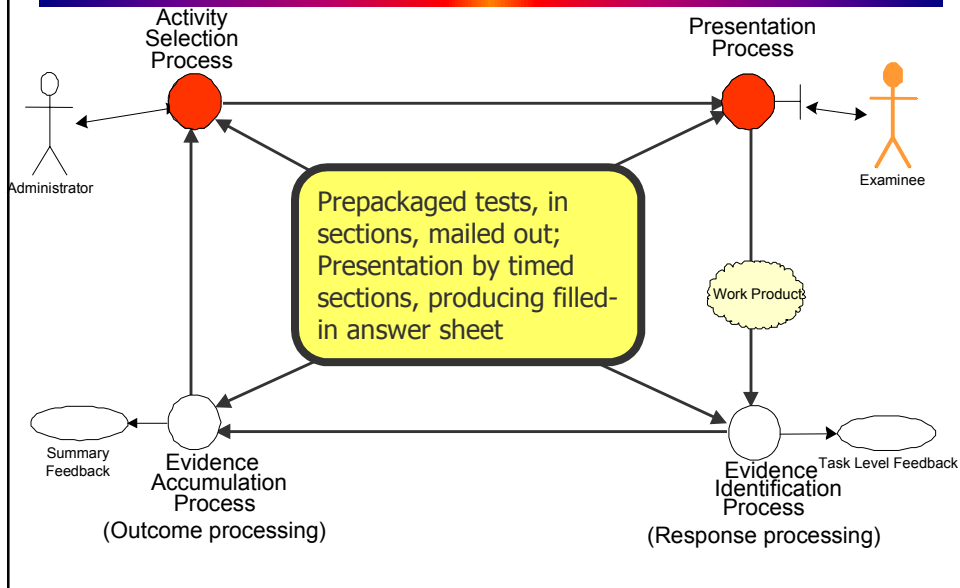


Four Ways of Using GRE Items

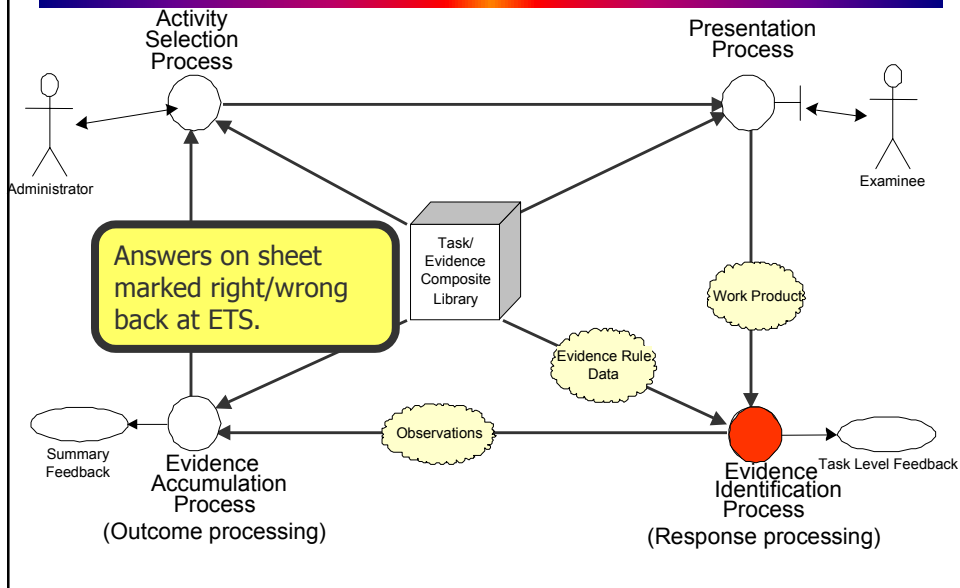
Note that in everyday language, “exactly the same items” are used in each case.

- 🌱 “Classic” Paper & Pencil GRE
- 🌱 Computerized-Adaptive Test
- 🌱 Interactive practice items on the Web
- 🌱 GRE on-line diagnostic service

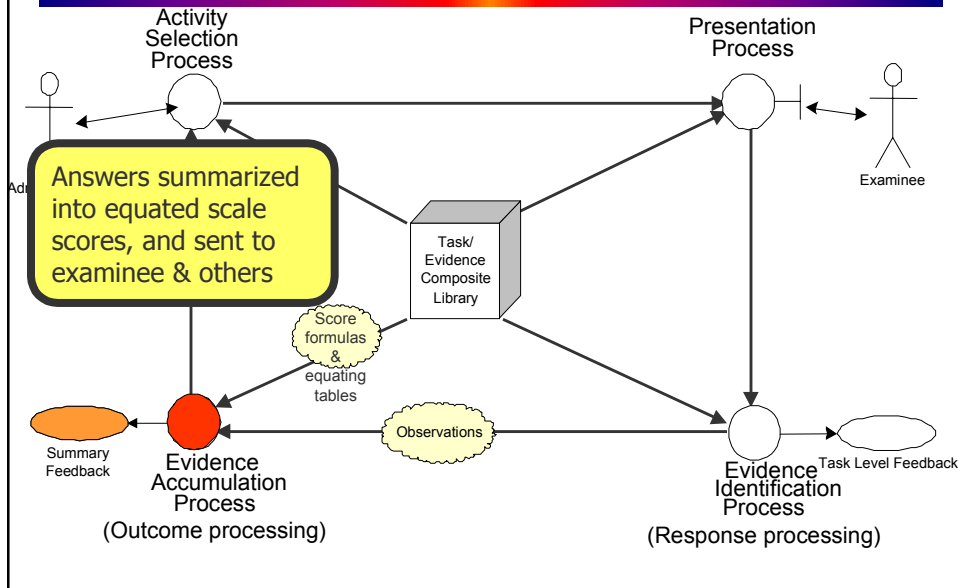
“Classic” Paper & Pencil GRE



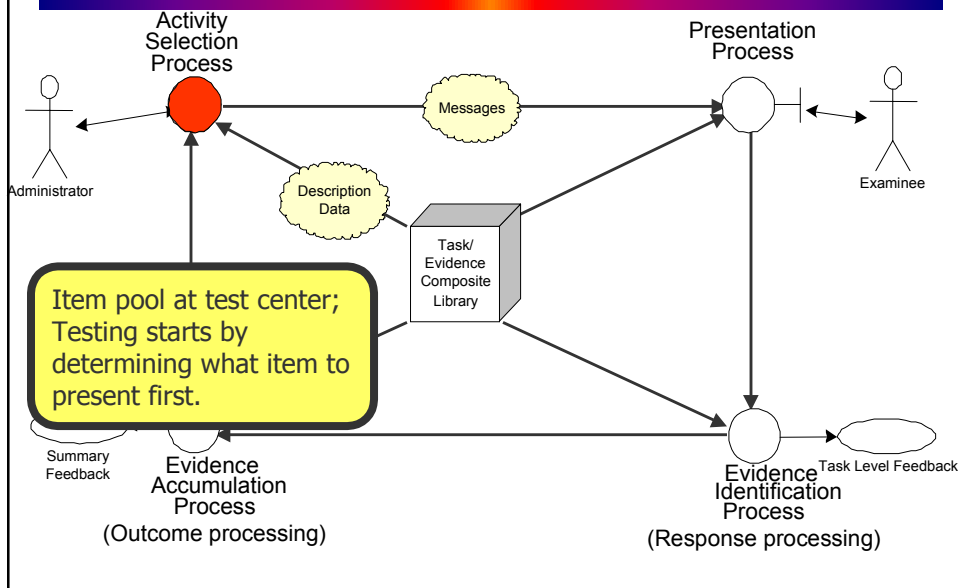
“Classic” Paper & Pencil GRE



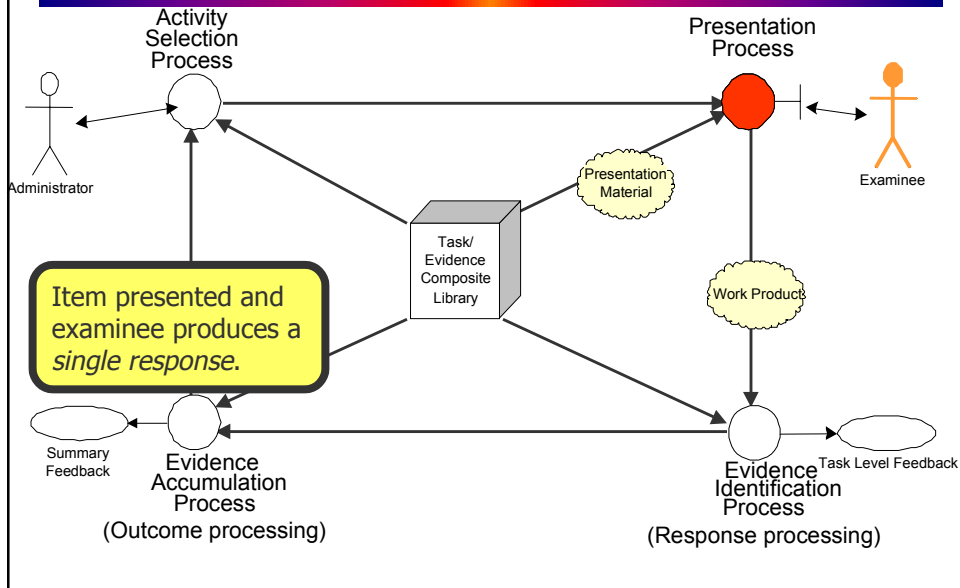
“Classic” Paper & Pencil GRE



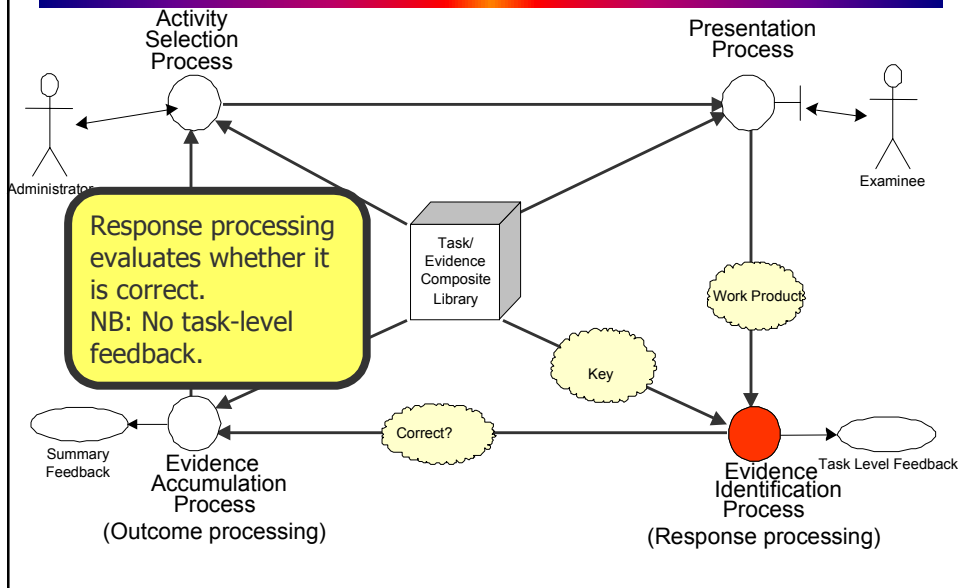
Computerized-Adaptive Test



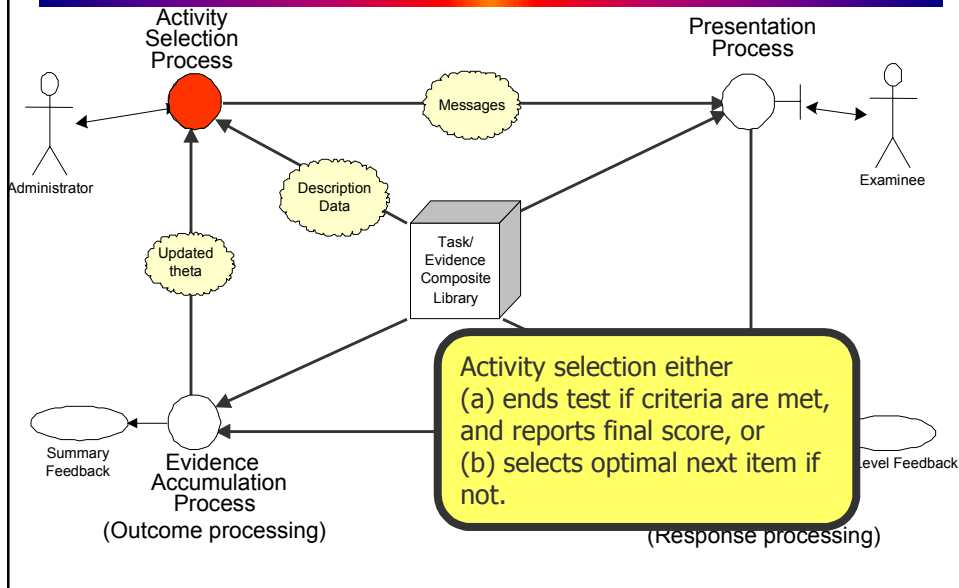
Computerized-Adaptive Test



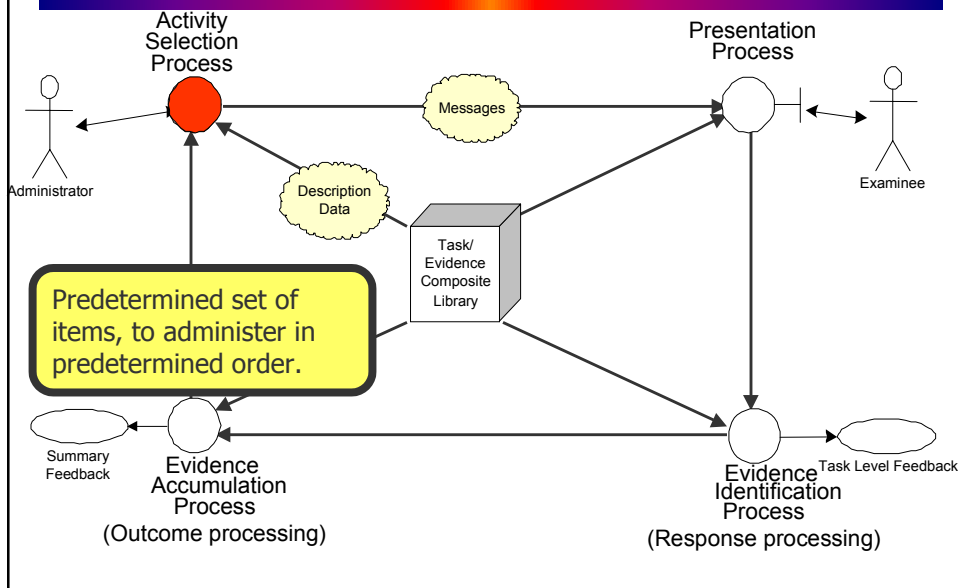
Computerized-Adaptive Test



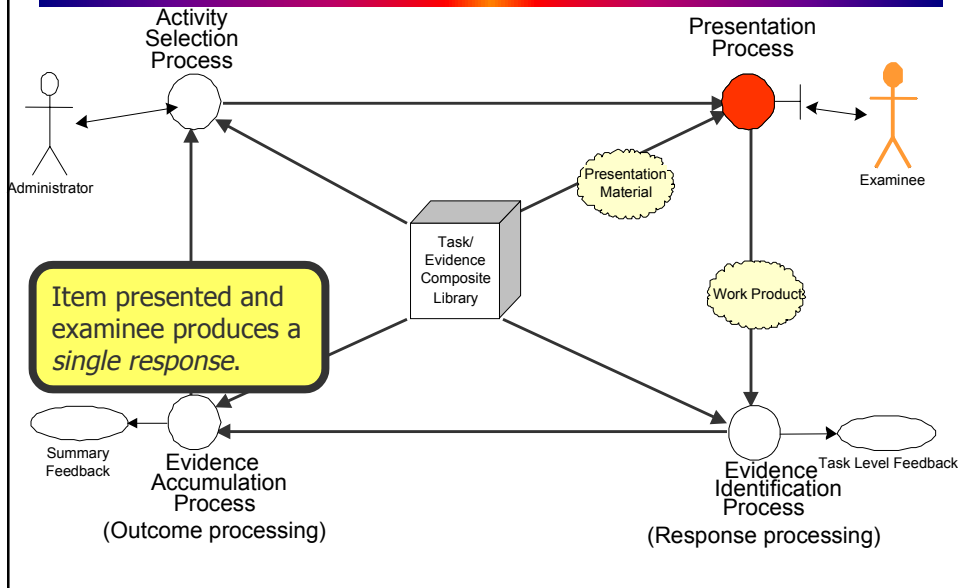
Computerized-Adaptive Test



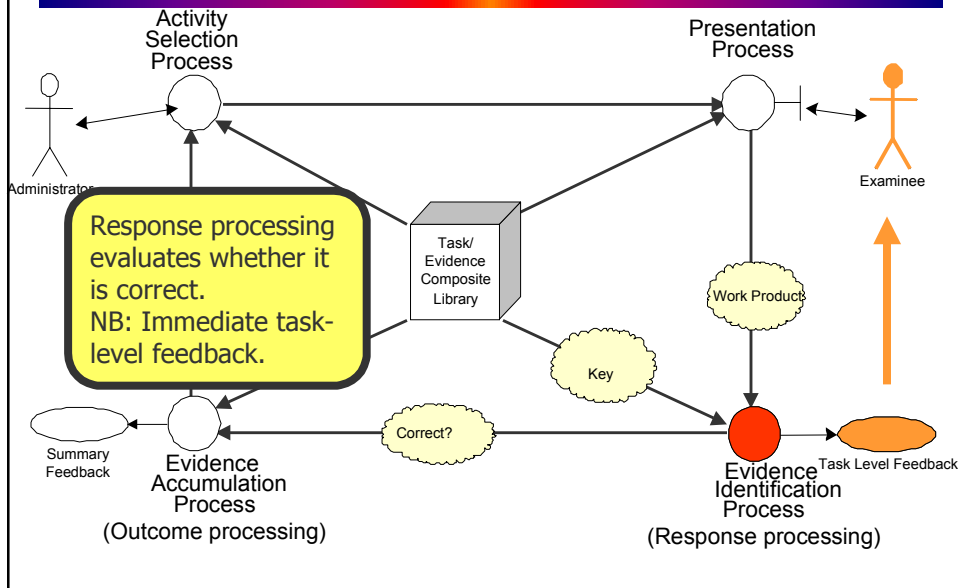
Interactive practice items on the Web



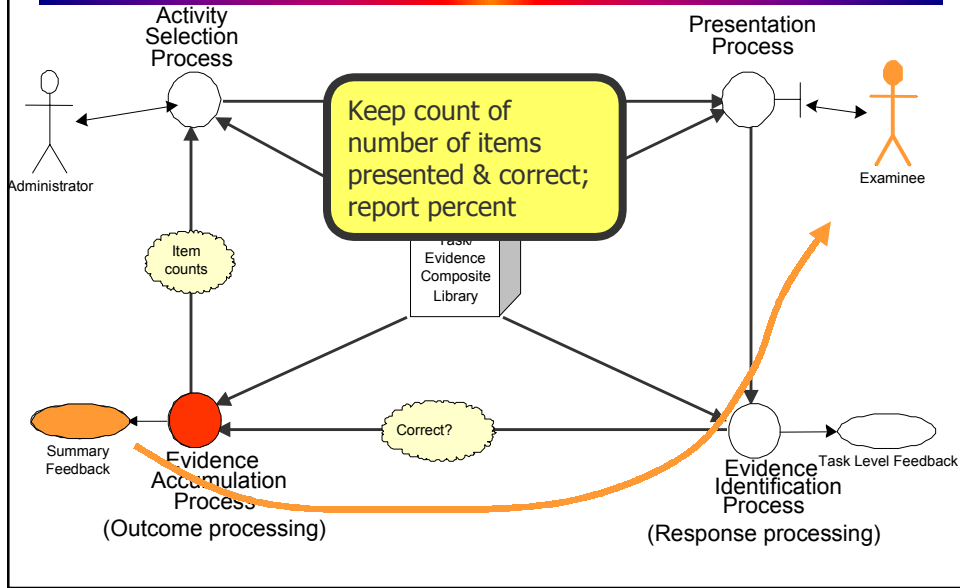
Interactive practice items on the Web



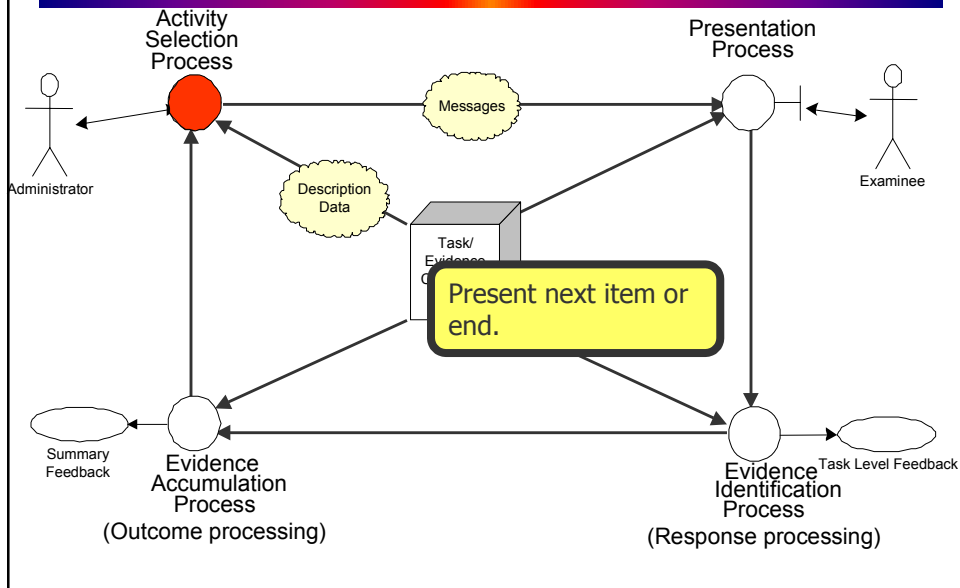
Interactive practice items on the Web



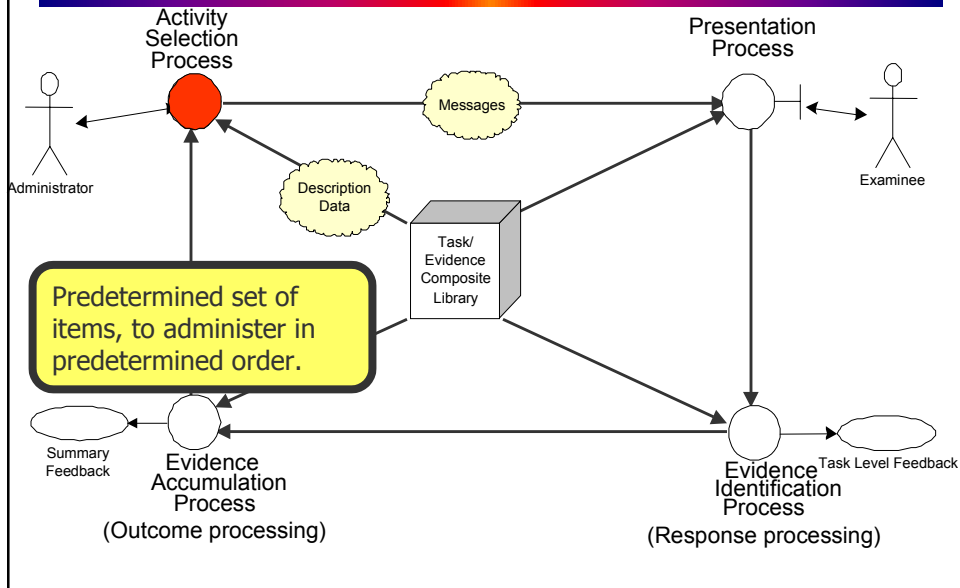
Interactive practice items on the Web



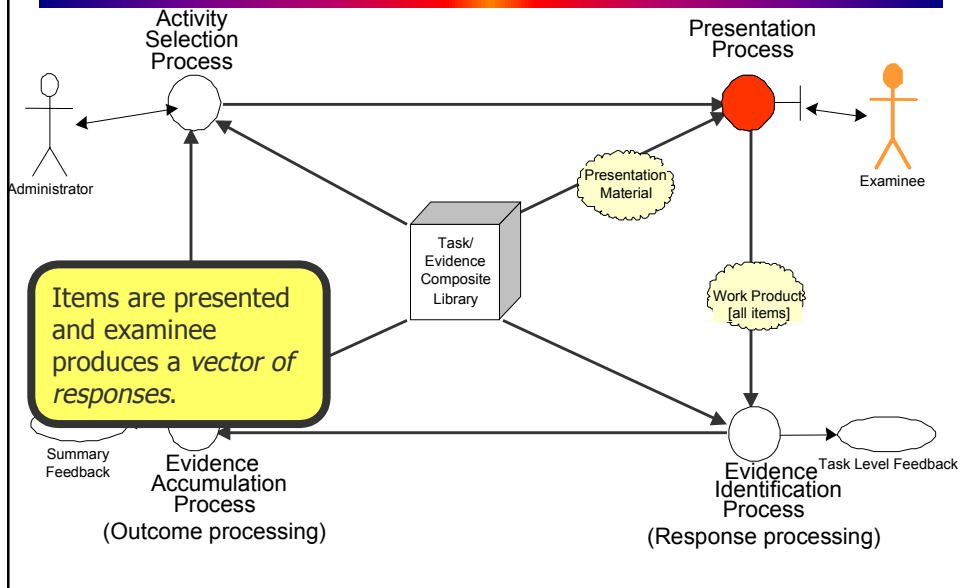
Interactive practice items on the Web



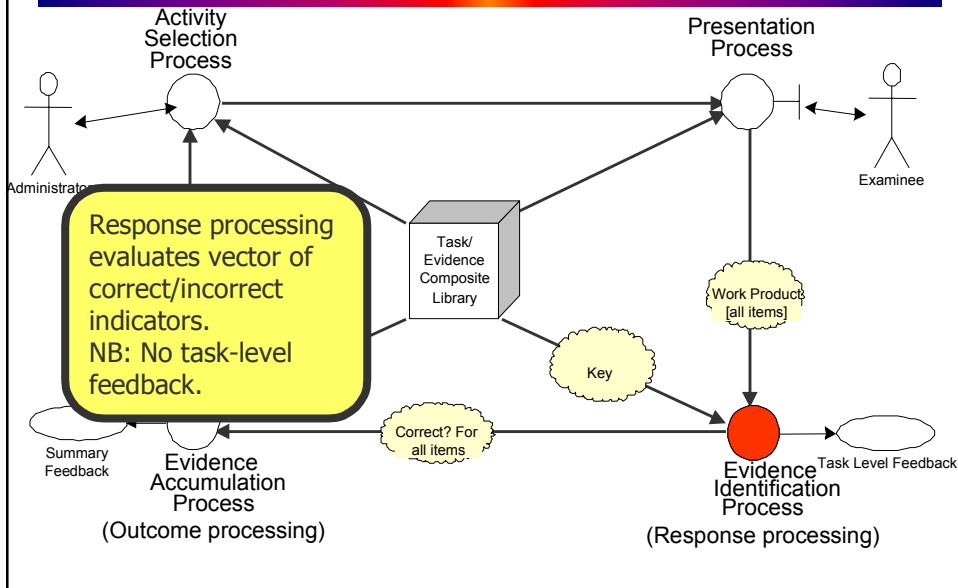
GRE on-line diagnostic service



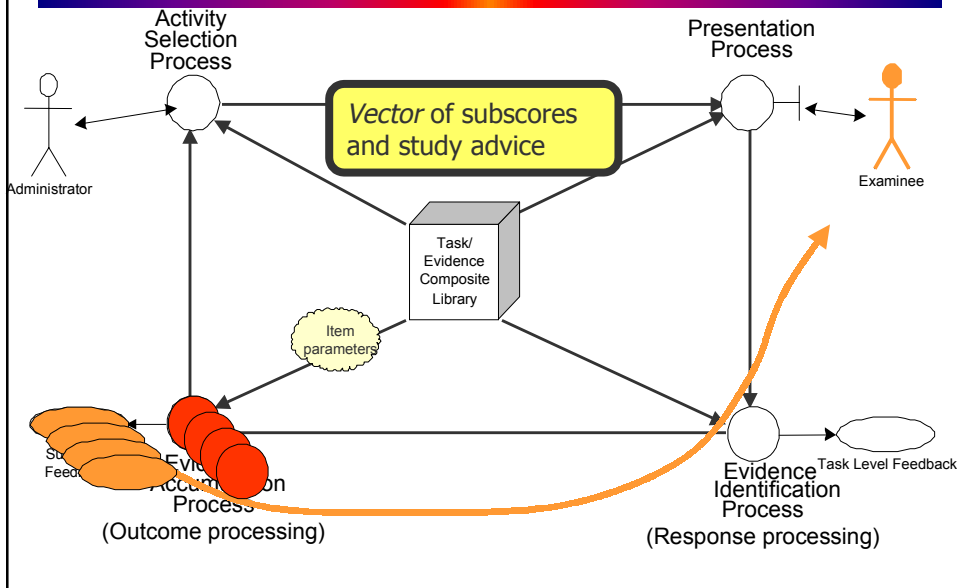
GRE on-line diagnostic service



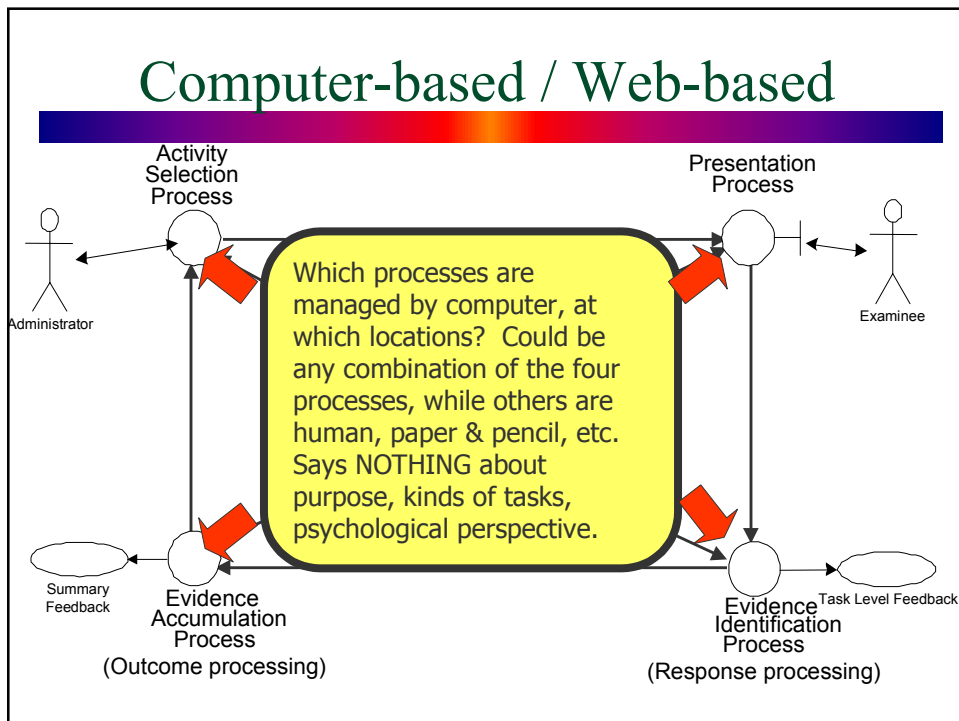
GRE on-line diagnostic service



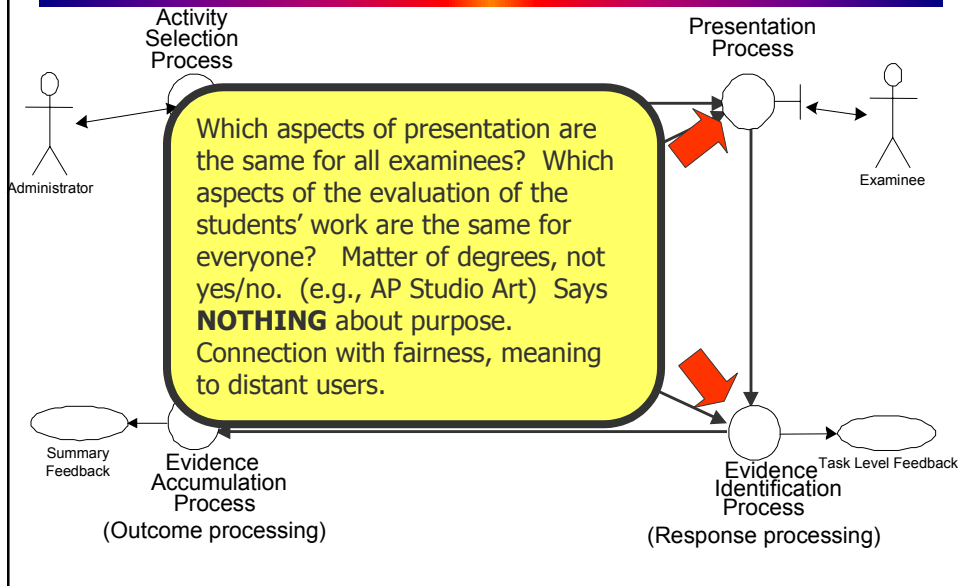
GRE on-line diagnostic service



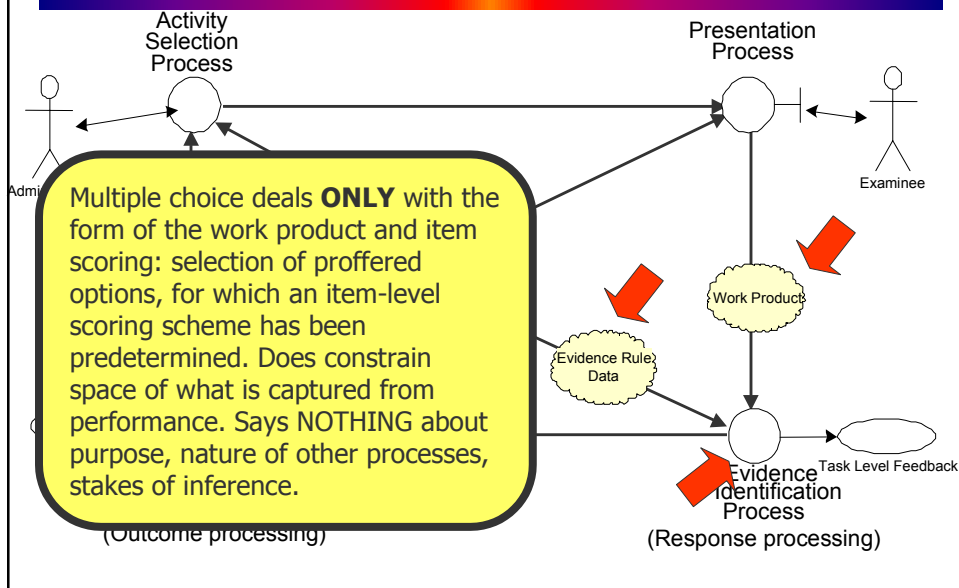
Now, what do those phrases mean?



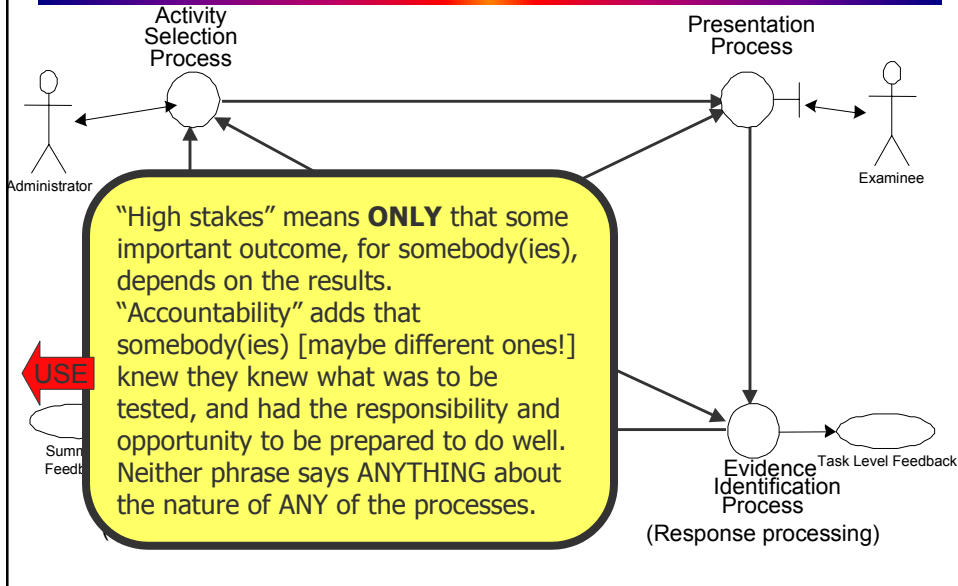
Standardized Test



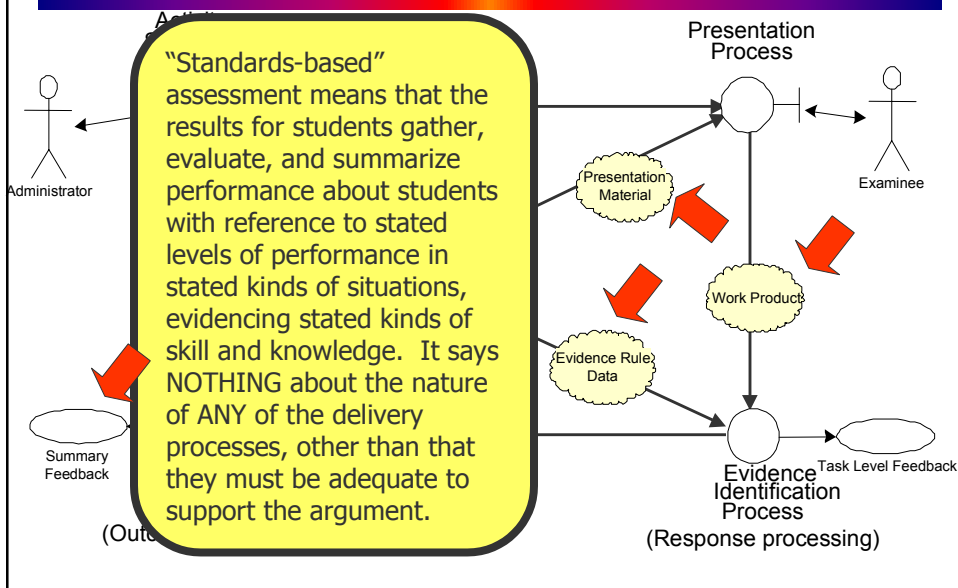
Multiple-choice Testing



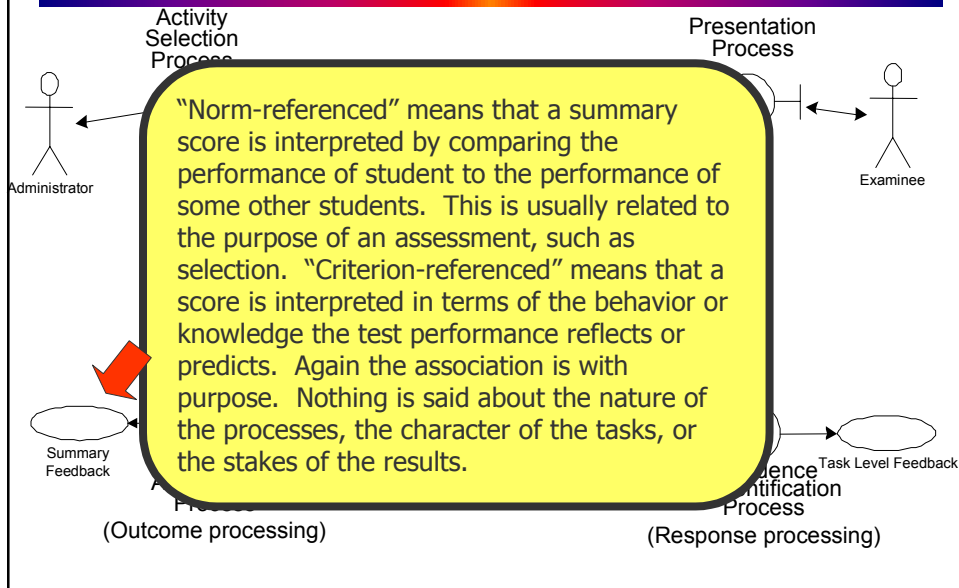
High-Stakes / Accountability



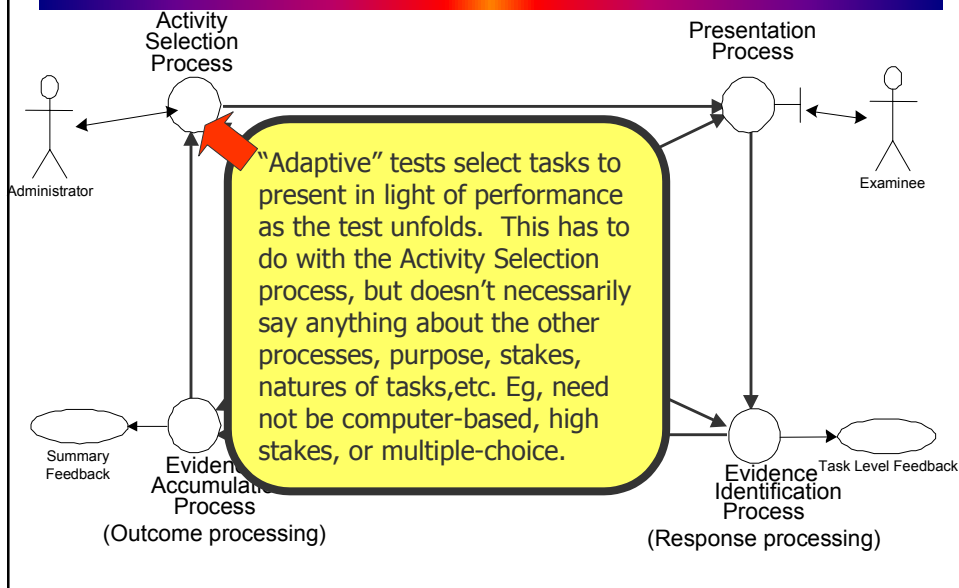
Standards-Based



Norm- vs. Criterion-Referenced



Adaptive test



Conclusion

- Progress in assessment requires progress in language.
- ETS/IMS four-process delivery system provides a schema and a vocabulary to describe rigorously the processes in assessment, and the information that must be passed among them.