

# The Four-Process Assessment Delivery System

---

Russell G. Almond, Linda S. Steinberg,  
and Robert J. Mislevy

Educational Testing Service  
June 6, 2000

This work was supported by the Research Division and the New Product Development  
Offices of Educational Testing Service.

# A quote from Don Norman, Part 1:

---

Infrastructure is usually thought to be dull. Tedious. Few people wish to think about it until it is necessary, which is then often too late. Once established, it is expensive and often difficult to change. Moreover, infrastructures require standardization; they're too expensive and restrictive to allow multiple infrastructures to coexist, too important to society to allow the monetary interests of one company or industry to determine the underlying infrastructure for everyone.

...

## A quote from Don Norman, Part 2:

---

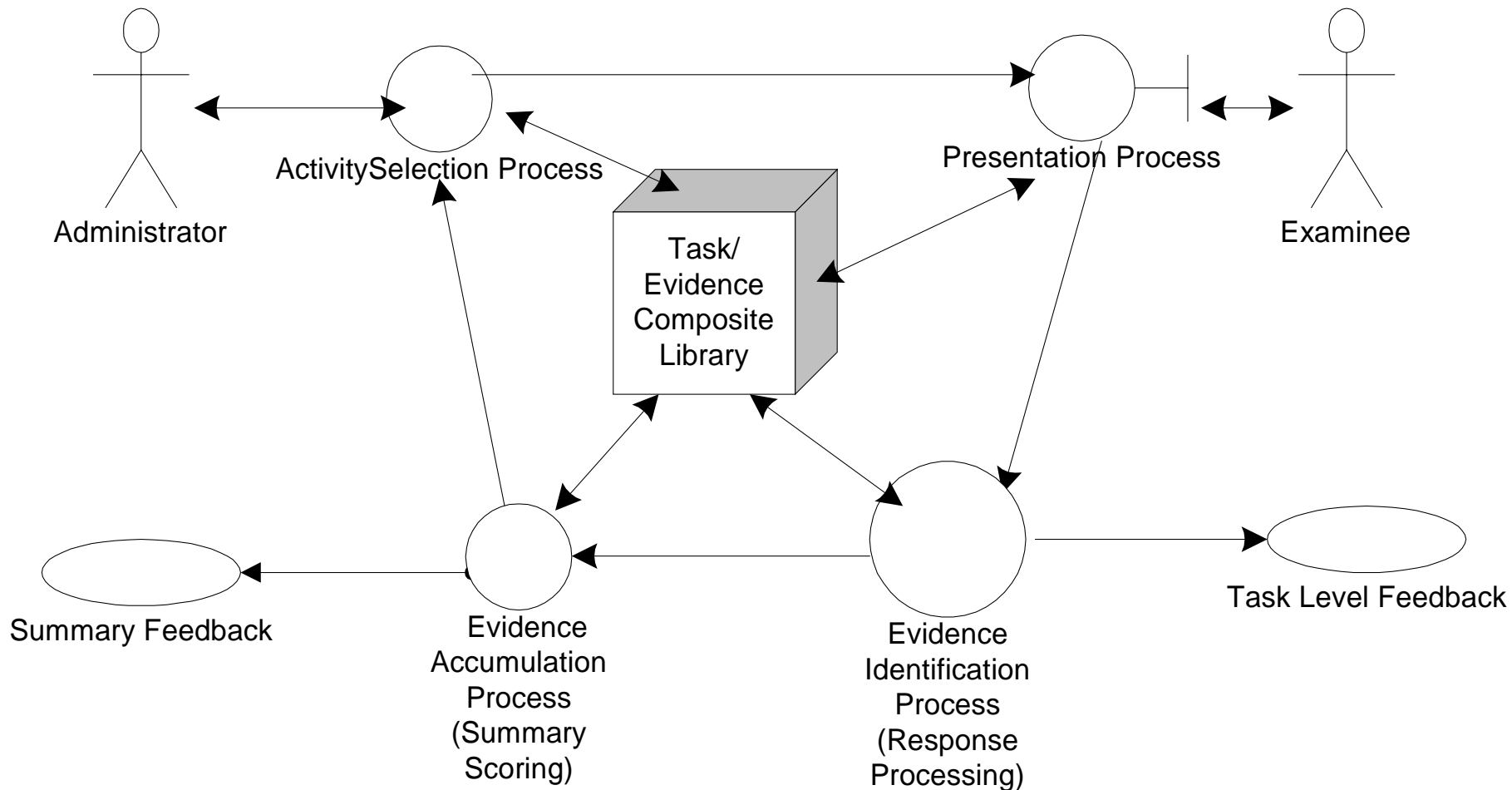
Probably the most important lesson for the development of information appliance industry is the importance of establishing an open, universal standard for exchanging information. If only we can establish world-wide standards for the sharing of information, then the particular infrastructure used within each appliance becomes irrelevant. Each appliance can use whatever best fits its needs. Each company can select whatever infrastructure makes most sense to its operations. Once the information exchange is standardized, nothing else matters.

(Norman, 1998, pp.. 132-133)

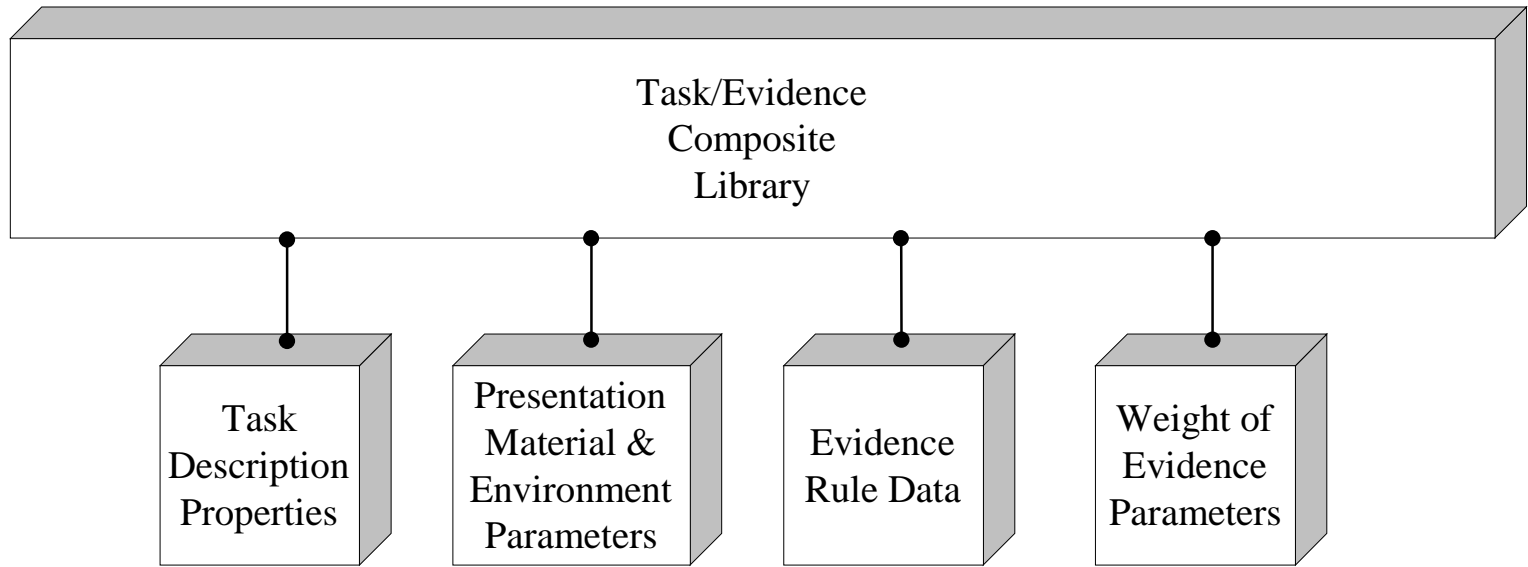
# Origins of the Four-Process Model

---

- Evidence-centered assessment design
  - How to implement assessment designs based on ECD model & design process
  - Maximize re-use of assessment pieces and processes
- The IMS project
  - Initial IMS work focused on *format* of exchange data
  - ETS “white paper” on IMS information model extends to information *about what, and why*

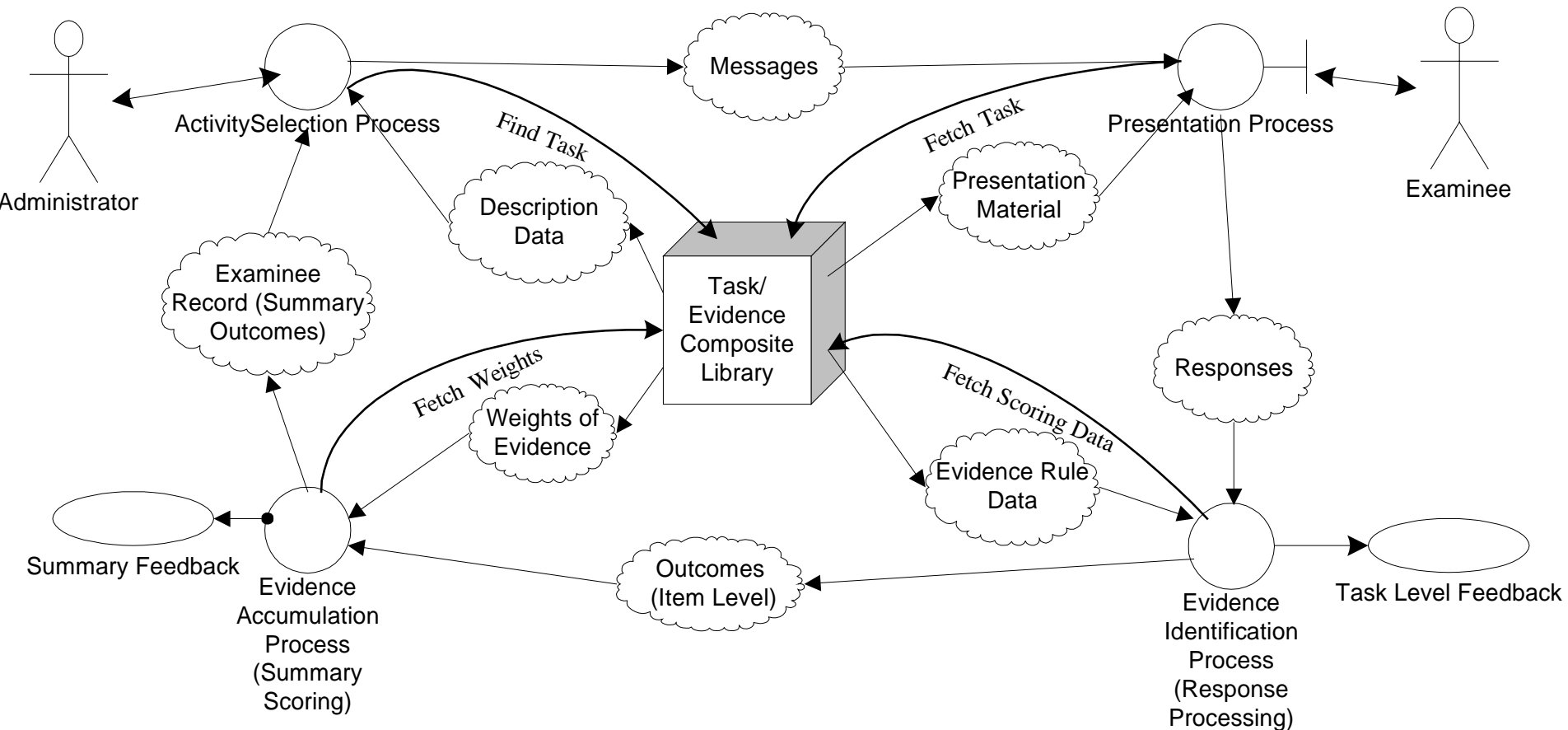


This figure shows the four principle processes in the assessment cycle. The Activity Selection Process selects a task (or other activity) and instructs the Presentation Process to display it. When the examinee has finished interacting with the item, the Presentation Process sends the results (a collection of responses) to the Evidence Identification Process for item-level Response Processing. This process identifies key outcomes of the task and passes them to the Evidence Accumulation Process (Section or Assessment level scoring), which updates the Examinee record. The Activity Selection then makes a decision about what to do next based on whatever criteria are appropriate, including, for example, tasks already completed or current beliefs about the examinee.



The Task/Evidence Composite Library contains information about each task needed to...

- Present it
- Allow the student to interact with it (including available aides such as a spell-checker or calculator)
- Capture the student's response (i.e., specs for the work product)
- Extract meaningful features from the work product
- Evaluate meaningful features in terms of observable variables (i.e., task-level scoring)
- Integrate evidence into an accumulating student record (i.e., for test- or section-level scoring; e.g., item parameters, item weights)



This is a more detailed view of the assessment cycle. Here we expand the picture to show the data objects taken from the task/evidence library and passed around the cycle.

# Examples based on the GRE

---

Exactly the same domain of tasks in all cases, but different delivery modes & processes to suit different purposes.

- The classic paper & pencil GRE
- GRE-CAT
- Print version of “real GRE” with explanations
- GRE on-line diagnostic service