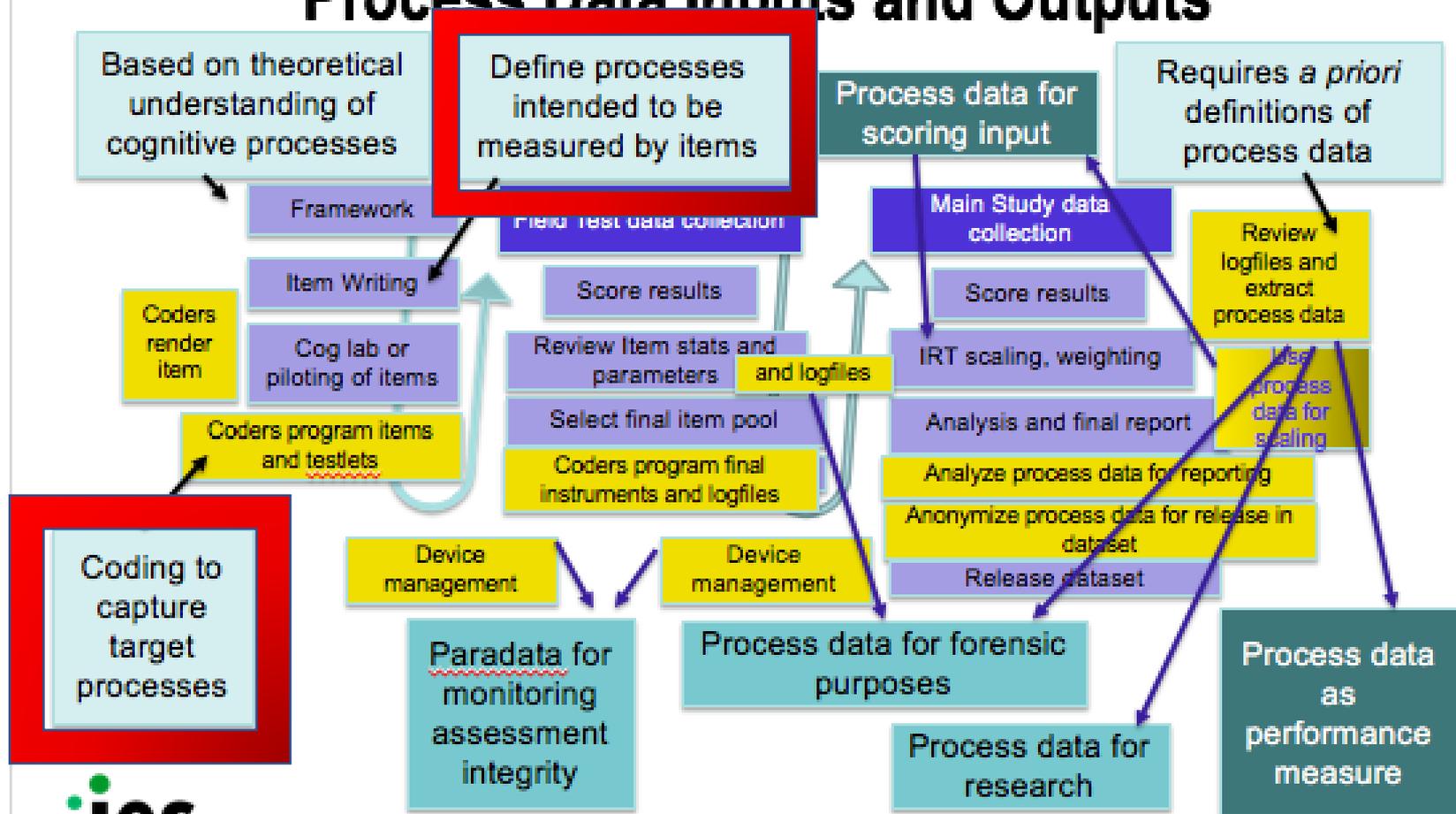# Opportunity versus Challenge:
Exploring Usage of Log-File and Process Data in International Large Scale Assessments

Retrospective Points

1. Items we are analyzing to mine process data were not written to support process data analysis, which complicates efforts to do.

2. If we are to have process data to test hypotheses, items need to be set up to capture appropriate processes.

If we are to have process data to test hypotheses, items need to be set up to capture appropriate steps.


Process Data Inputs and Outputs

Based on theoretical understanding of cognitive processes

Define processes intended to be measured by items

Process data for scoring input

Requires *a priori* definitions of process data

Framework

Field Test data collection

Main Study data collection

Review logfiles and extract process data

Item Writing

Score results

Score results

Coders render item

Cog lab or piloting of items

Review Item stats and parameters

and logfiles

IRT scaling, weighting

Use process data for scaling

Coders program items and testlets

Select final item pool

Analysis and final report

Coders program final instruments and logfiles

Analyze process data for reporting

Anonymize process data for release in dataset

Coding to capture target processes

Device management

Device management

Release dataset

Paradata for monitoring assessment integrity

Process data for forensic purposes

Process data as performance measure

Process data for research

25

Visualization of NAEP reading patterns from sampled logfiles

TIME LAPSE 0:00 ANIMATION SPEEDS 2X 10X 50X ▶ ❚❚

PAGE 1 PAGE 2 PAGE 3 PAGE 4 PAGE 5

Each sampled student represented by a blue dot.

Start
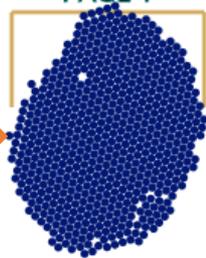
Pages of text represented by "roofs" indicating which page being looked at.

Ten test questions represented by "buckets"

1 2 3 4 5 6 7 8 9 10

QUESTIONS

REVIEW

Retrospective Points

1. Items we are analyzing to mine process data were not written to support process data analysis, which complicates efforts to do.

2. If we are to have process data to test hypotheses, items need to be set up to capture appropriate processes.

3. "Time on task" is complicated to use as a variable (and total time needs to be distinguished from time on specific processes)

Time on task

Loris: Measure of effort – proxy indicator developed from time data.

Britt (Quiwe): Time on task can be generalized and abstracted from PSTRE items but no difference between time spent on task leading to a correct answer and those leading to a wrong answer.

Reinaldo: Can't discriminate success and failure based on time

Kentaro: Response time data interact with
1. Student proficiency/skills
2. Assessment domain
3. Language of assessment
4. Cognitive demand of items (e.g., low skilled will read at the same speed but high skilled readers will take more time reading on harder items)

Retrospective Points

4. Process data allow us to move beyond simple correct/incorrect.

5. Distinguishing between different types of "correct" and "incorrect" responses (based on how achieved/how failed) provides valuable insight on student performance.

6. Important to distinguish between "information generation" (taking correct steps to arrive at the correct conclusions) and "information interpretation" (translating conclusions into the correct response).

# Research into understanding respondent behaviors and cognitive strategies

Luis and Frank: research to improve teaching and learning with specific information on how different students think/perform

Sam, Britt, and Reinaldo: research to better understand factors that distinguish high- and low-performers, or expert from novice strategies

Loris: research to better understand the relationship of motivation and performance.

Research into understanding respondent behaviors and cognitive strategies
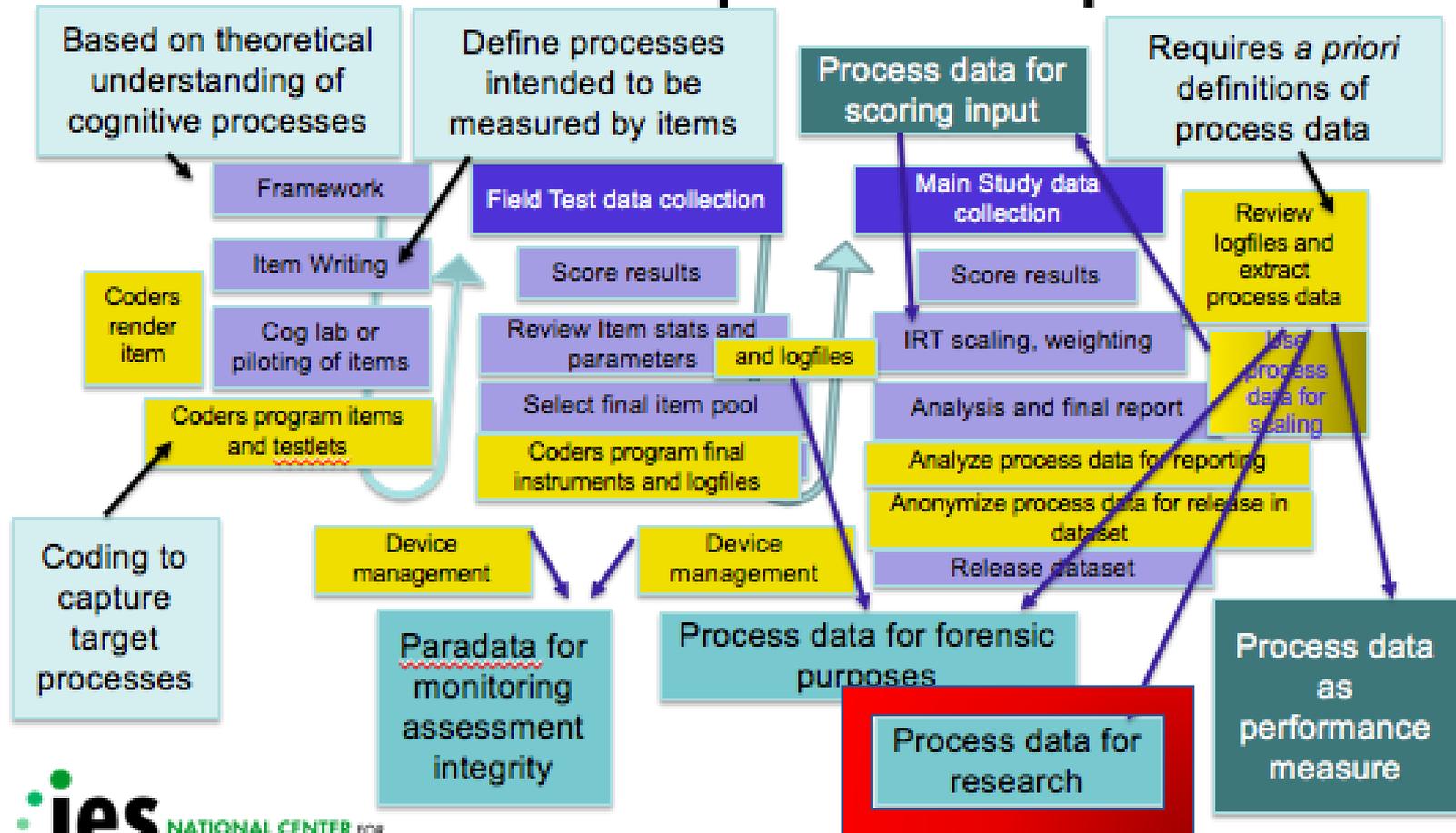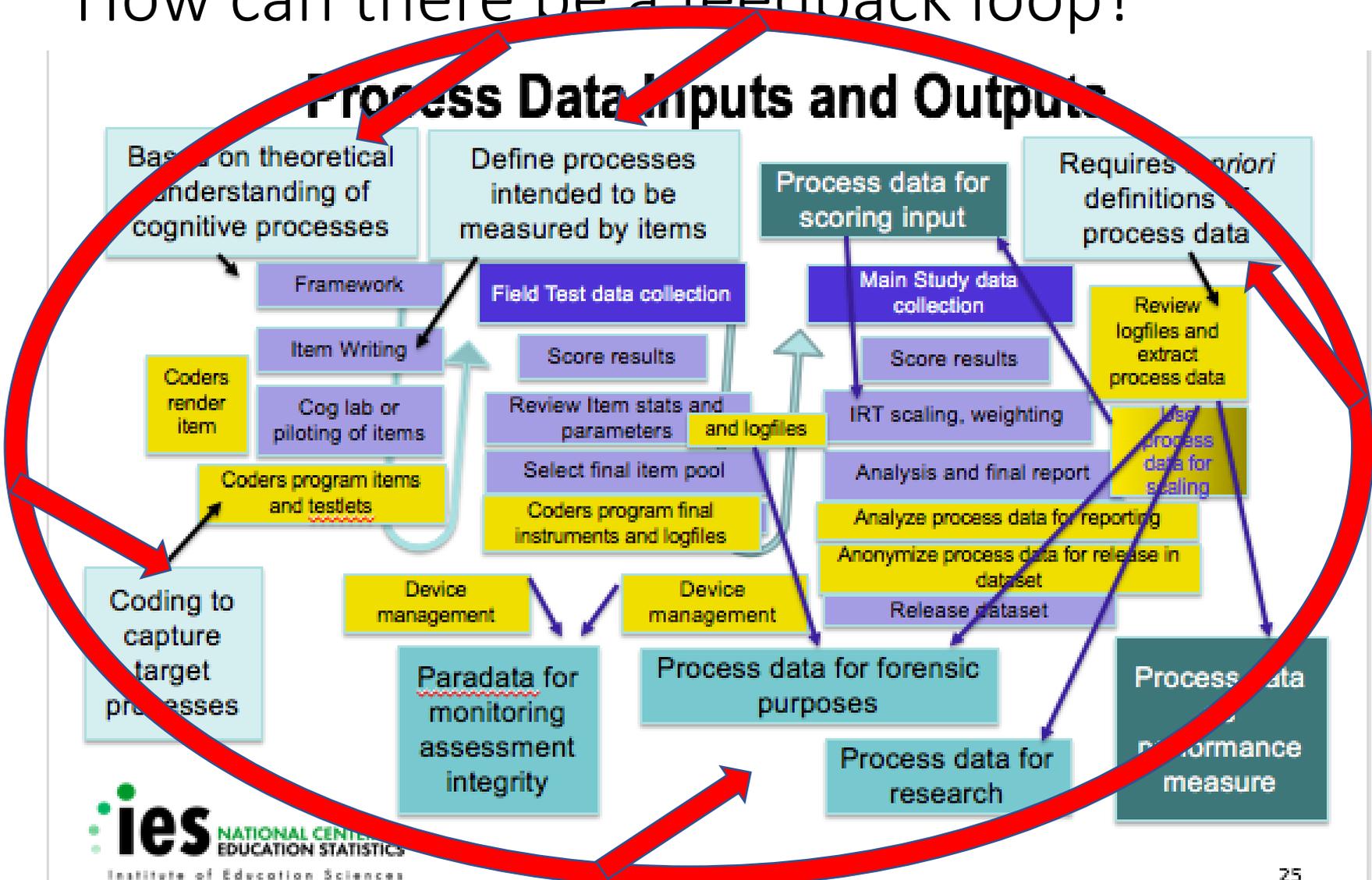
Retrospective Points

7. Iterative feedback loop essential between design and analysis to improve quality of process data and research

How can there be a feedback loop?

Process Data Inputs and Outputs

Retrospective Points

7. Iterative feedback loop essential between design and analysis to improve quality of process data and research

8. Not clear how large-scale assessment systems can be modified to incorporate feedback

Inappropriate uses of process data

For example,

- Overgeneralizing from one item to all items, or one process to many processes
- Concluding that strategies associated with higher performance are the strategies that all students should be taught
- Making classroom and formative assessments turn on process data in such a way that students lose unstructured opportunities to try out new ways of thinking and doing
- Ethically questionable uses of student responses and identification of students

Purpose of assessments

- Who can get the answer correct – sorting function
- Where do students go wrong? – pedagogical function to improve teaching
- How do humans learn? – research function
- What is nature of knowledge, thinking, and understanding? – philosophical question

Main takeaway

Ultimately process data are a proxy for cognitive processes