Evaluating the Evaluators: How do we bring it all together and begin?

Panel Discussion

Reidy Interactive Lecture Series

September 14, 2012

Center for Assessment

Pulling it All Together

- Yesterday we discussed fallible measures:
  - Test scores
  - Growth models/VAM
  - Observations
  - Surveys

- How does combining several imperfect measures lead to a better model?
  - Do we get a better understanding of teacher effectiveness than a single imperfect measure?
  - What are implications of combining fallible measures?
    - Devaluing one?
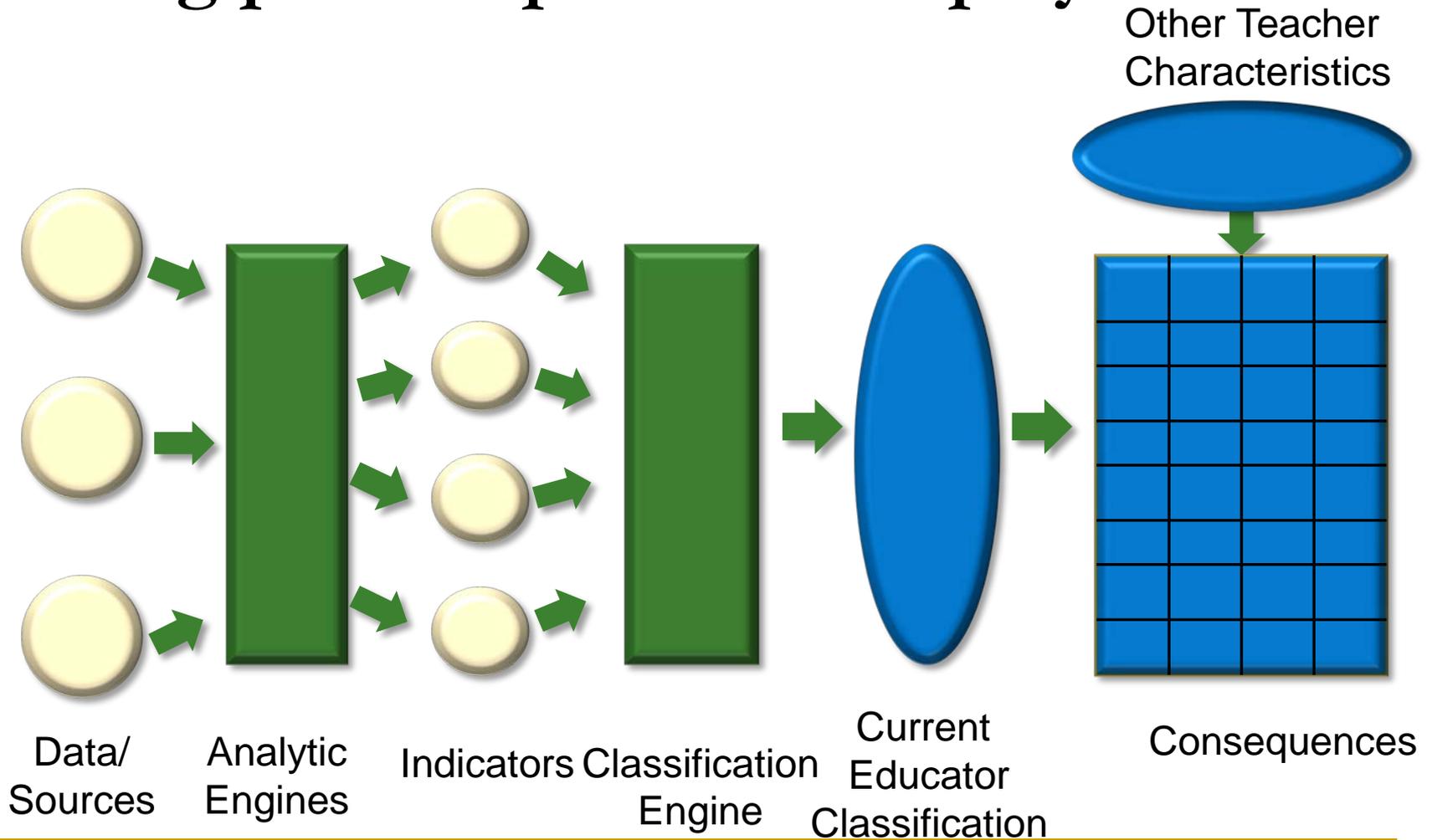   - Forcing one to match another?

Looking at Courtney's handout, there are lots of analyses. How do we prioritize where to start?
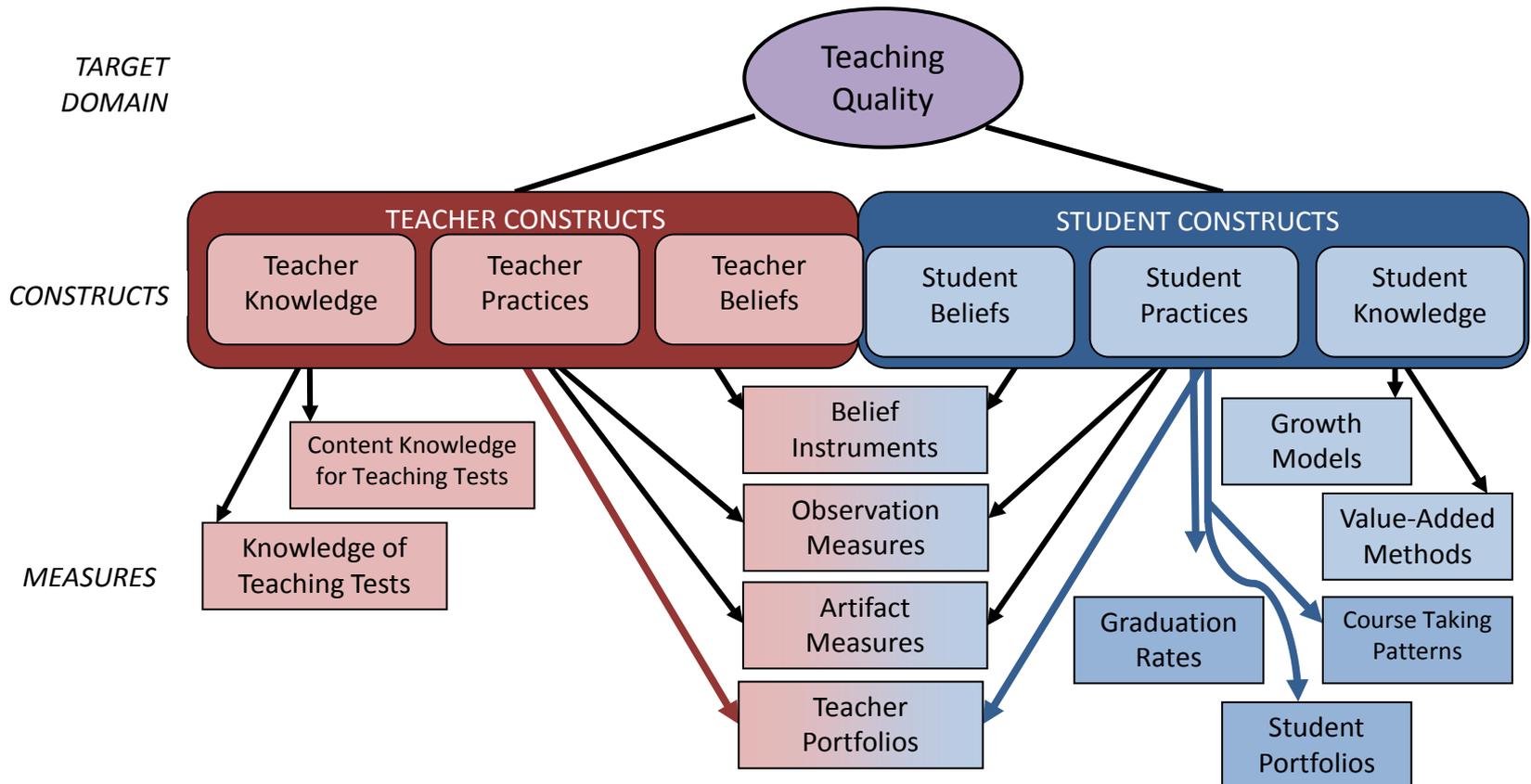
Adapted from Lorrie Shepard's work on prioritizing research studies

- ## What is critical to know?

- ## What do you really not know?
  - ❑ Give more weight to studies that provide new information rather than those that confirm what you already know

- ## What are the costs?
  - ❑ May choose the least expensive option, at least to start with

- ## Timing
  - ❑ Some studies will be more appropriate at certain points in the evaluation cycle
  - ❑ High-stakes decisions are being made immediately—where can we get the biggest bang for our buck?

Center for Assessment

Where should we start and does the starting point depend on the player?

Other Teacher Characteristics

Data/ Sources   Analytic Engines   Indicators   Classification Engine   Current Educator Classification   Consequences

If we buy that evaluating teaching quality is separate from evaluating teachers, how can we best fulfill the (legal?) obligation to collect the data and link it to individual teachers in a way that does the least harm?

TARGET DOMAIN

Teaching Quality

CONSTRUCTS

TEACHER CONSTRUCTS
- Teacher Knowledge
- Teacher Practices
- Teacher Beliefs

STUDENT CONSTRUCTS
- Student Beliefs
- Student Practices
- Student Knowledge

MEASURES

- Content Knowledge for Teaching Tests
- Knowledge of Teaching Tests
- Belief Instruments
- Observation Measures
- Artifact Measures
- Teacher Portfolios
- Growth Models
- Value-Added Methods
- Graduation Rates
- Course Taking Patterns
- Student Portfolios

Would we recognize an effective educator if one were standing right in front of us? If yes, would we run her over??
Or…how will we know when we have the unintended negative consequence of having good teachers exit the school systems? How can we set up a system to answer that question?

Center for Assessment

Other Questions

- Is an effective teacher effective will all students?
  - If not, how many?
  - How do we decide?
- Should all teachers be evaluated under the same model?
  - Elementary vs. Secondary?
  - Special populations vs. General education
  - State tested subject versus not

Center for Assessment