



INSIGHT
HAPPENS HERE

**Essential Policies and Practices
for Test Score Validation**

Andrew Wiley

Paper presented at the annual meeting for the
Association of Test Publishers
March 22, 2016 Orlando, FL

#atpconf

Assessment Lifecycle



Design

- Program Structure
- Validation Plan
- Content Specifications

Develop

- Content Production
- Content Review
- Field Testing

Deliver

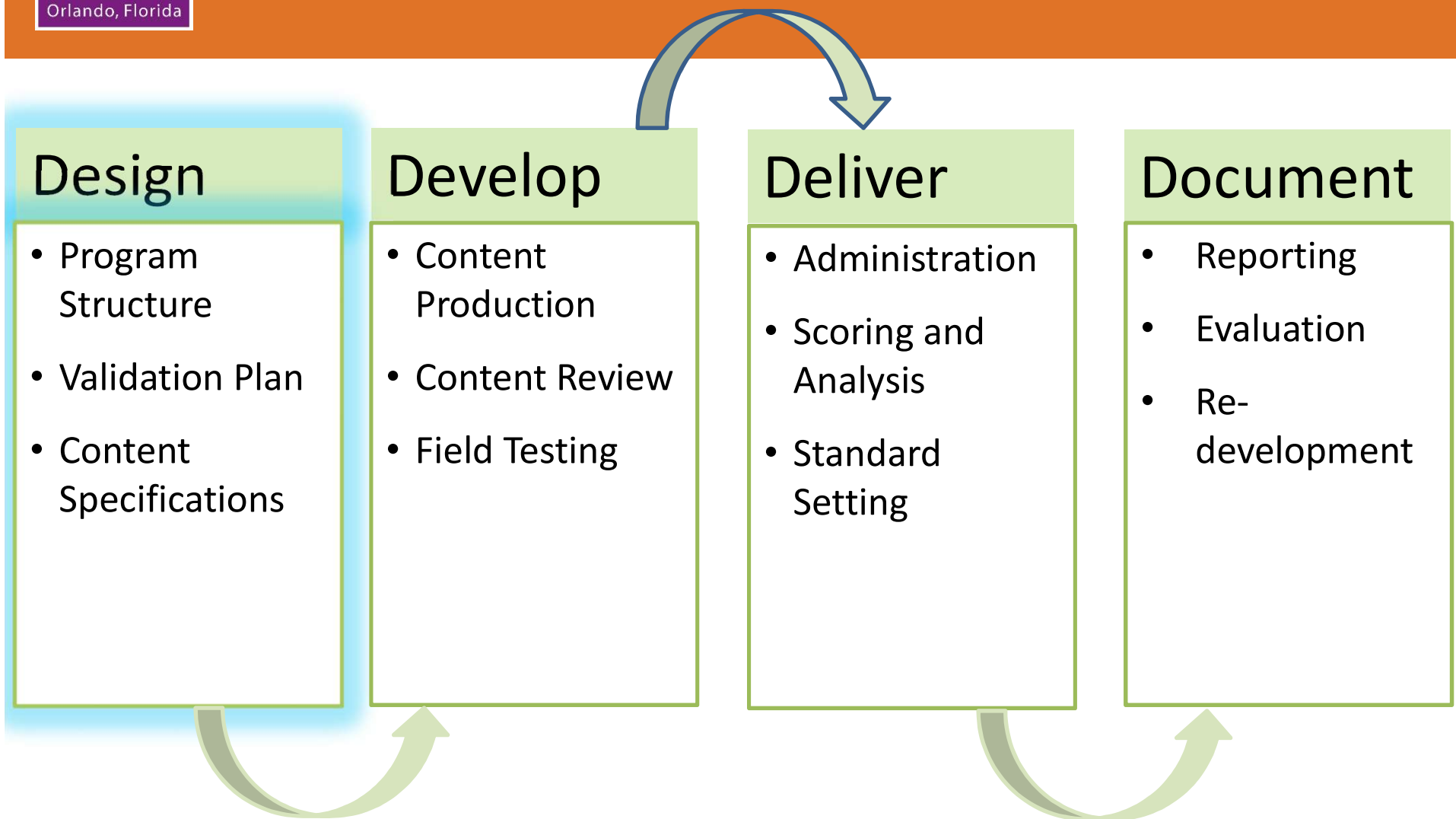
- Administration
- Scoring and Analysis
- Standard Setting

Document

- Reporting
- Evaluation
- Re-development



Assessment Lifecycle



- The validity argument starts with a statement that explains the claims that are intended to be made based upon the test scores
 - This student is college and career ready
 - This student is more/less qualified than other students for admissions into our school
 - This candidate has the knowledge, skills and abilities necessary to be certified in a given profession

Design: The key questions to ask

- Who was involved in the process?
- What was the process followed to complete this step?
- What were the outcomes of the process?

Design: The key questions to ask

- Who was involved in the process?
 - Did you use subject matter experts (i.e. professionals, teachers)
 - Did the people involve fully represent the population of people who will either take the test or be test user

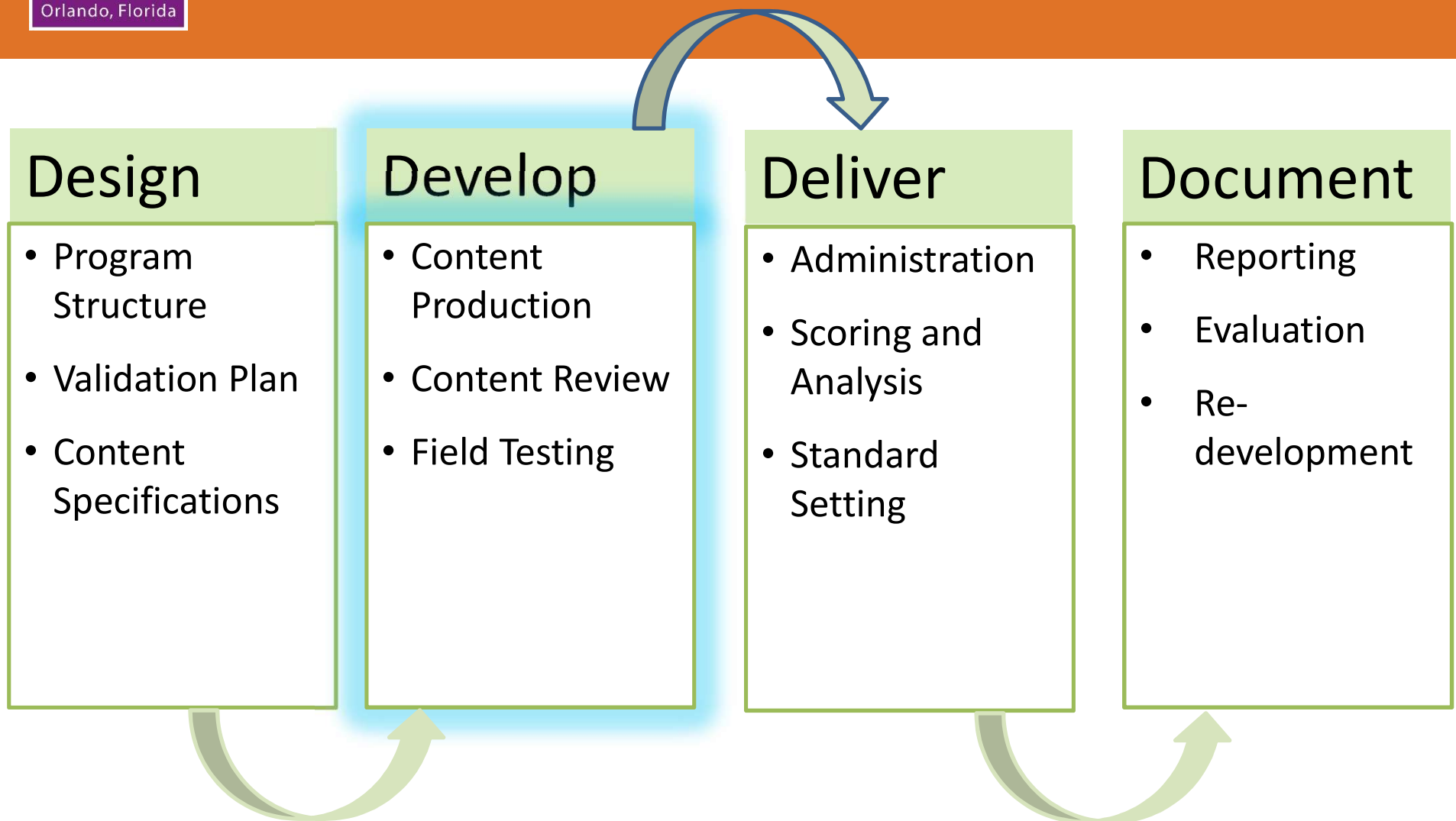
Design: The key questions to ask

- What was the process followed to complete this step?
 - Was there a systematic process for identifying the KSAs required to practice in a given profession?
 - Survey of current practices
 - Subject matter expert review

Design: The key questions to ask

- What were the outcomes of the process?
 - Is there a clear connection between the domain being measured on the test and the standards being assessed?
 - Is there a clear connection between the requirements for a person practicing in a given field and the content and domain on the test?

Assessment Lifecycle



Develop: The key questions to ask

- Who was involved in the process?
- What was the process followed to complete this step?
- What were the outcomes of the process?

Develop: The key questions to ask

- Who was involved in the process?
 - Where the people creating test items experts in the knowledge and domain being assessed?
 - Where the people involved familiar with the students or candidates who would be taking the given test?

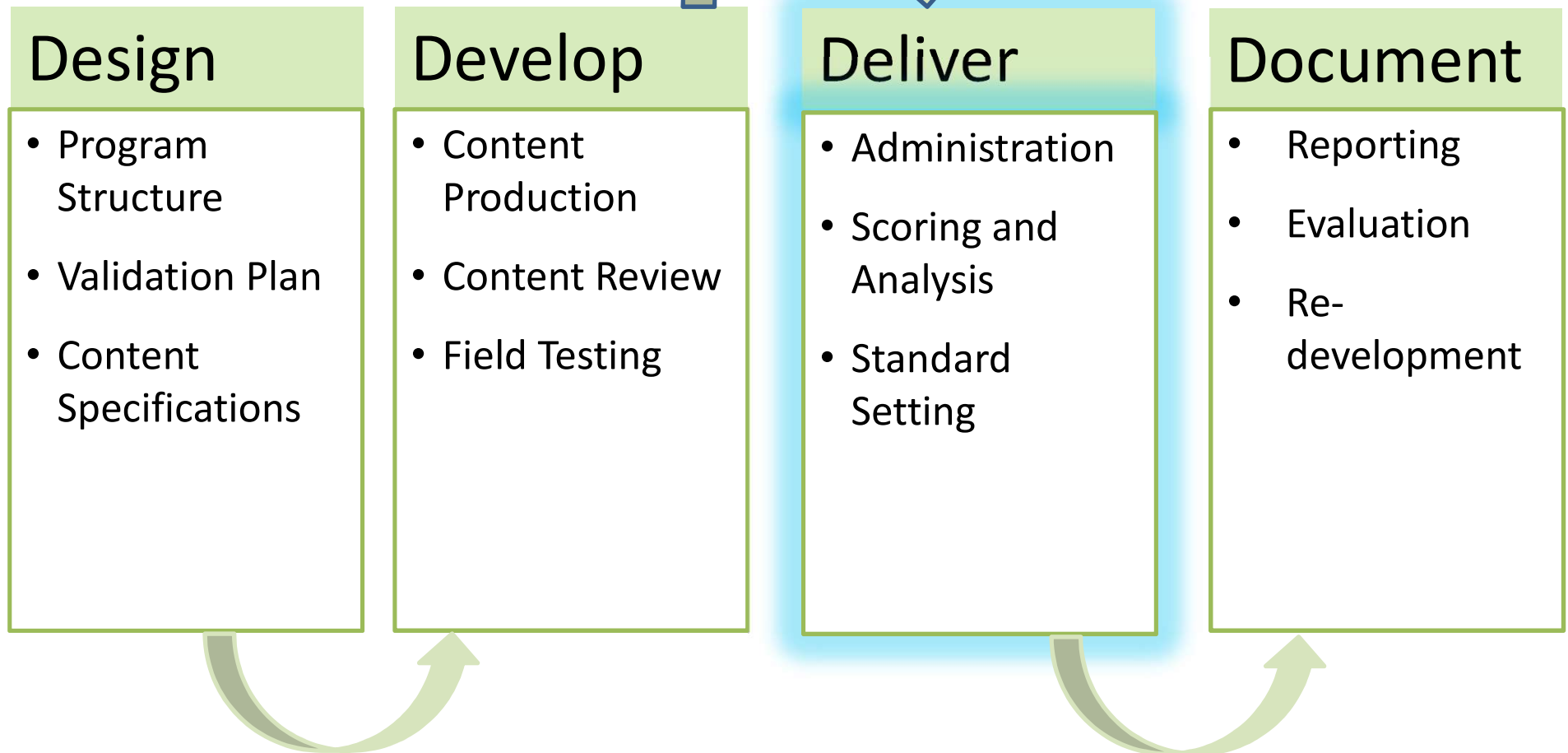
Develop: The key questions to ask

- What was the process followed to complete this step?
 - Was there any process designed to ensure candidates were answering items as intended (response process)?
 - Where the item writers trained in the process of writing good test items?
 - What was the review process when items were developed?
 - How were the items piloted prior to use on an operational test form?

Develop: The key questions to ask

- What were the outcomes of the process?
 - Were the items and tests structured in a way that is consistent with the construct being measured?

Assessment Lifecycle



Deliver: Standard Setting

■ Evaluation criterion

- Procedural evidence
- Internal evidence
- External evidence

Deliver: Standard Setting

■ Evaluation criterion – Procedural

- Qualifications of the panelists
- Clarity of the goals/tasks assigned to the standard setting panelists
- Data collection procedures
- Panelist confidence in their ratings and in the outcome of the process

Deliver: Standard Setting

■ Evaluation criterion – Internal

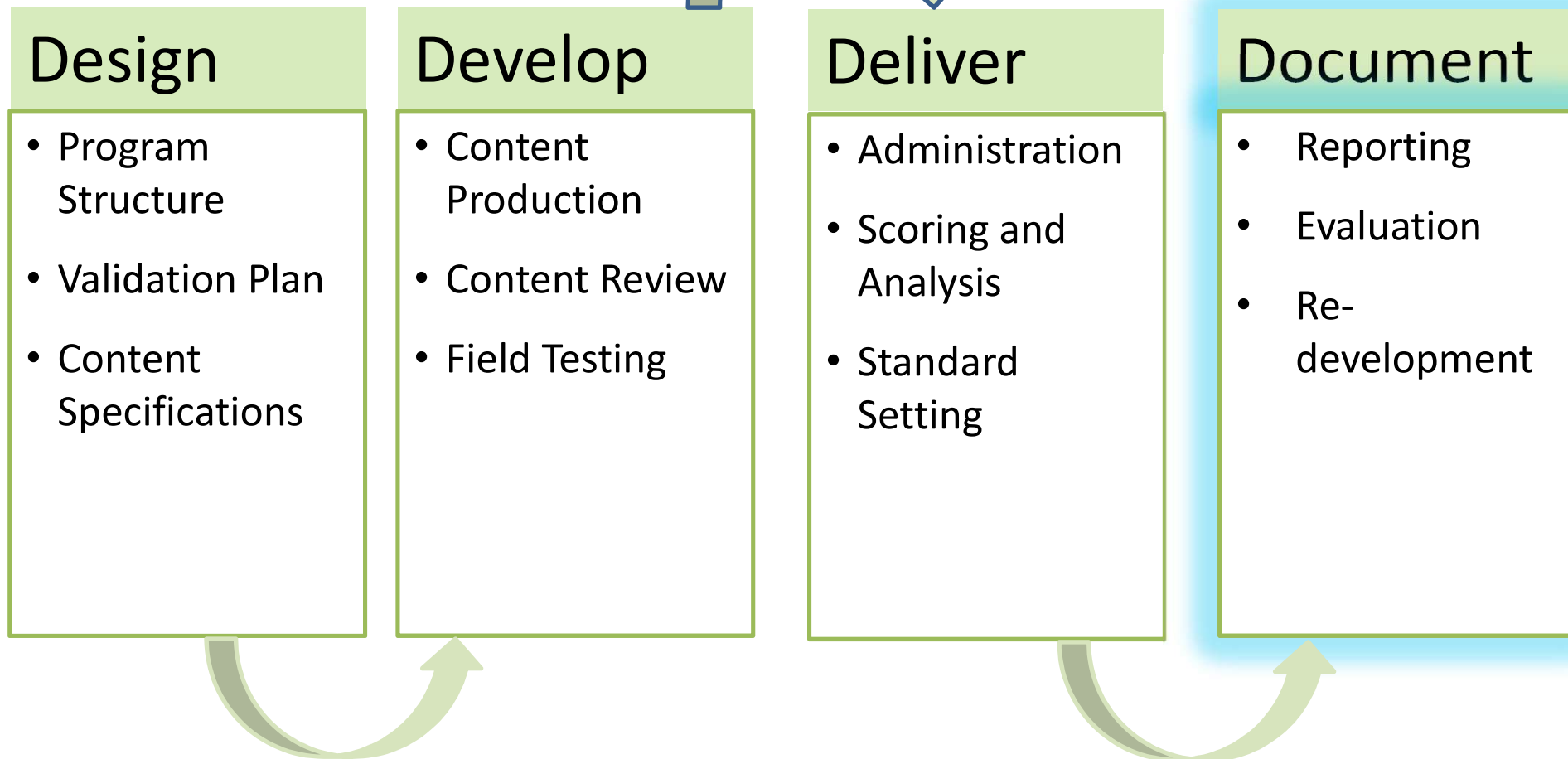
- Consistency of ratings across panelists
- Multiple rounds of ratings, with decreased variability in later rounds
- Independent panels

Deliver: Standard Setting

■ Evaluation criterion – External

- Consistency of passing rates with other similar measures
- Evaluation of performance across multiple groups
- Opportunity to review the *reasonableness* of the results in relation to program goals

Assessment Lifecycle





Document

- Relationships with other variables
- Reporting: Consequential validity



INSIGHT

HAPPENS HERE

Thank You!

Andrew Wiley
Awiley999@gmail.com

#atpconf