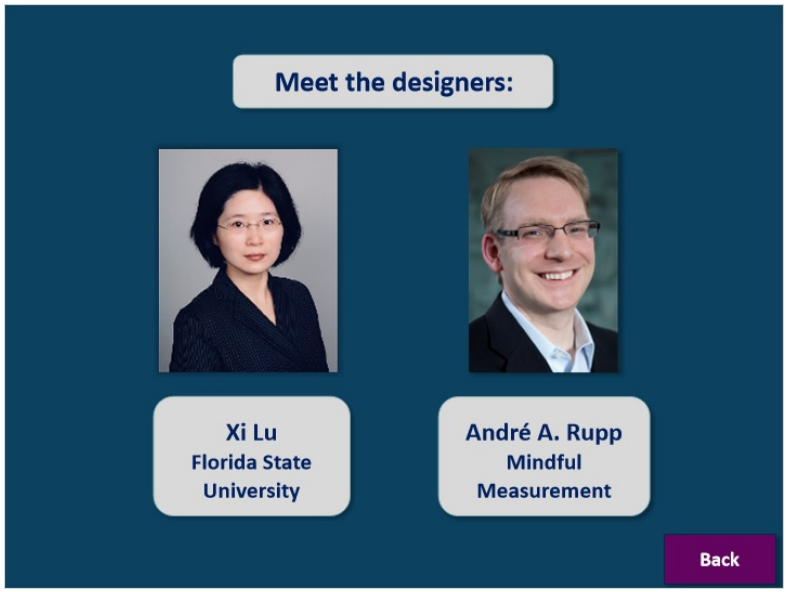
DM21 SLIDES (Results Reporting, Version 1.0)

1. Module Overview

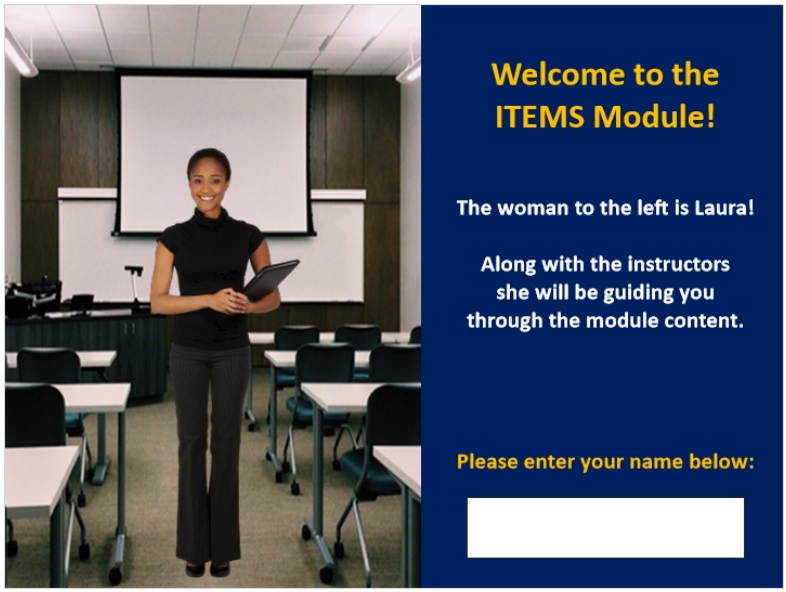
1.1 Module Cover (START)

1.2 Instructors

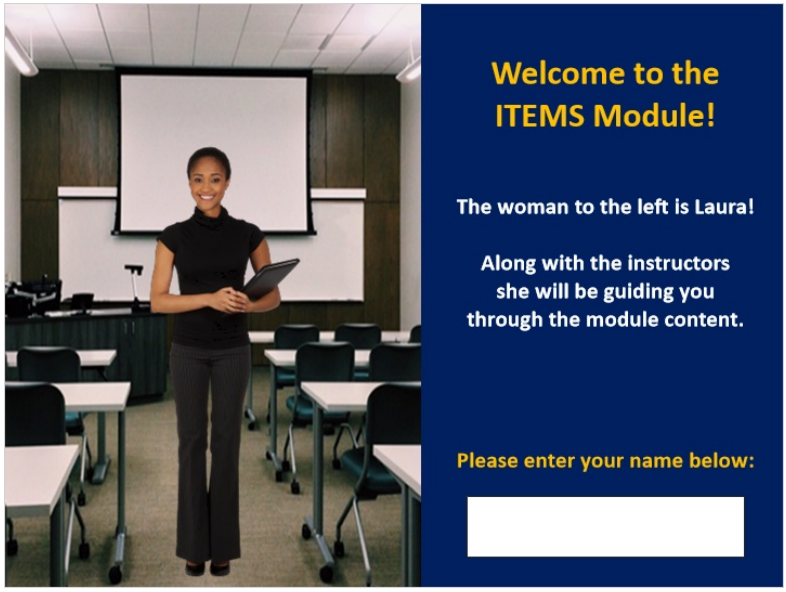
1.3 Designers



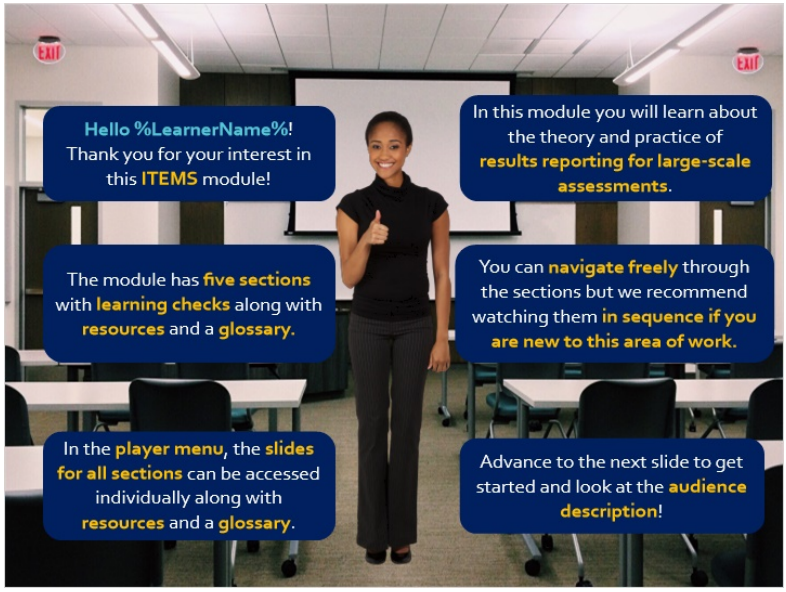
1.4 Welcome



Untitled Layer 1 (Slide Layer)



1.5 Overview



1.6 Target Audience

Target Audience

Anyone who would like a gentle but methodologically sound introduction to this topic:

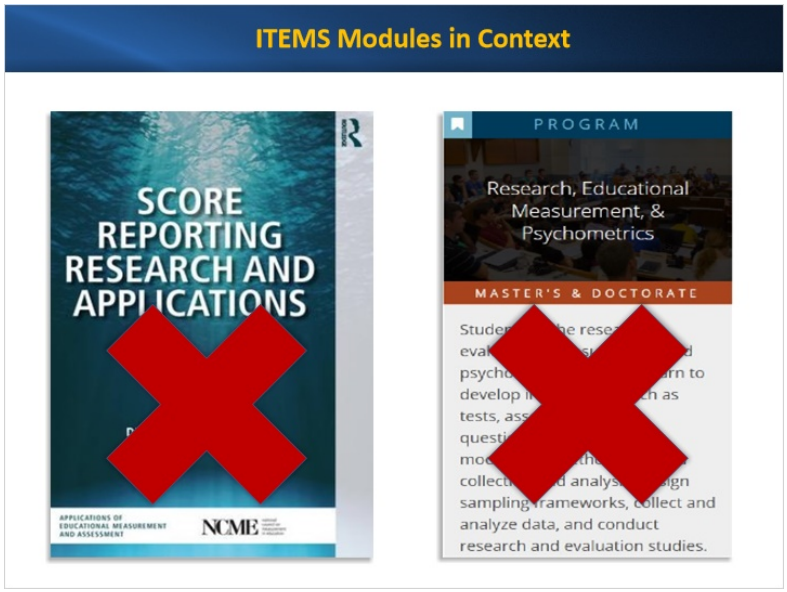
- graduate students and faculty in Master's, Ph.D., or certificate programs
- psychometricians and other measurement professionals
- data scientists / analysts
- research assistants or research scientists
- technical project directors
- assessment developers

However, we hope that you find the information in this module useful no matter what your official title or role in an organization is!

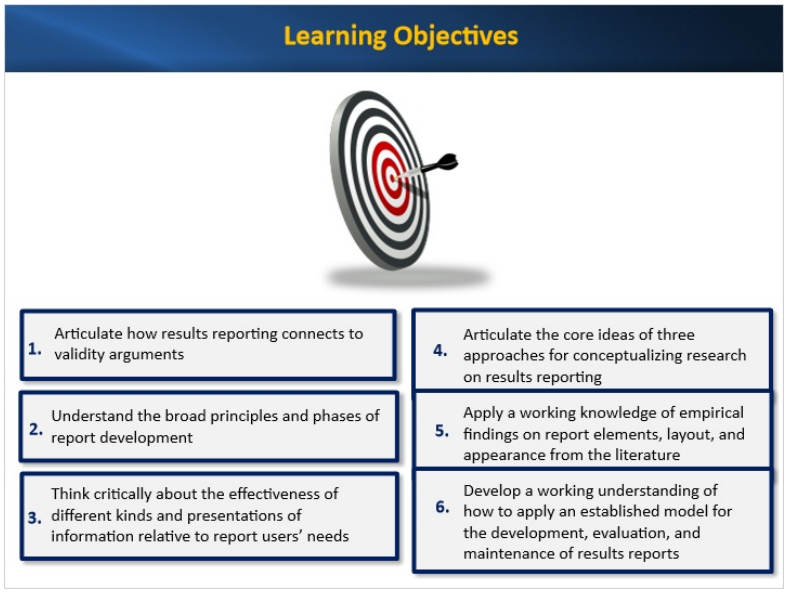
1.7 Expecations (I)

Let's discuss expectations....

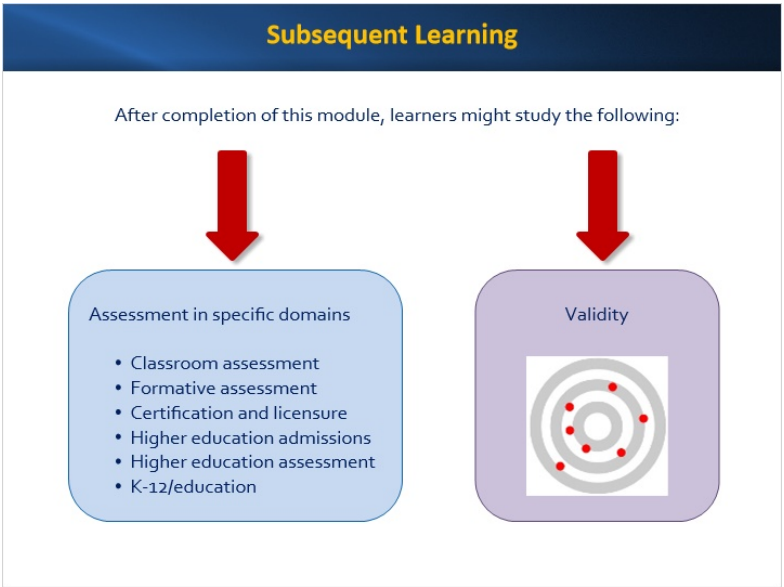
1.8 Expectations (II)



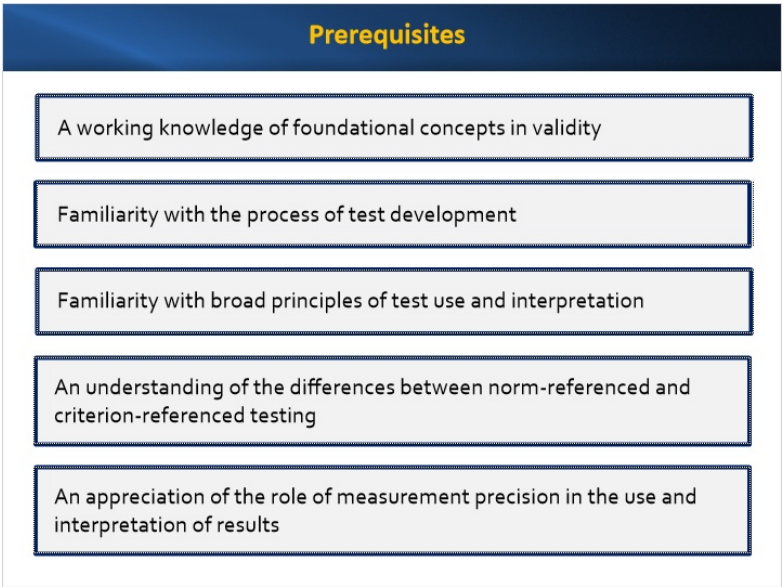
1.9 Learning Objectives



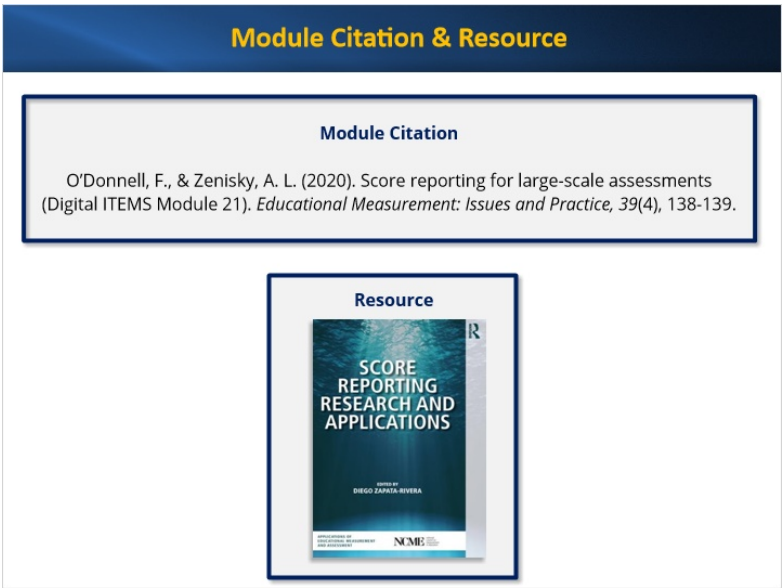
1.10 Subsequent Learning



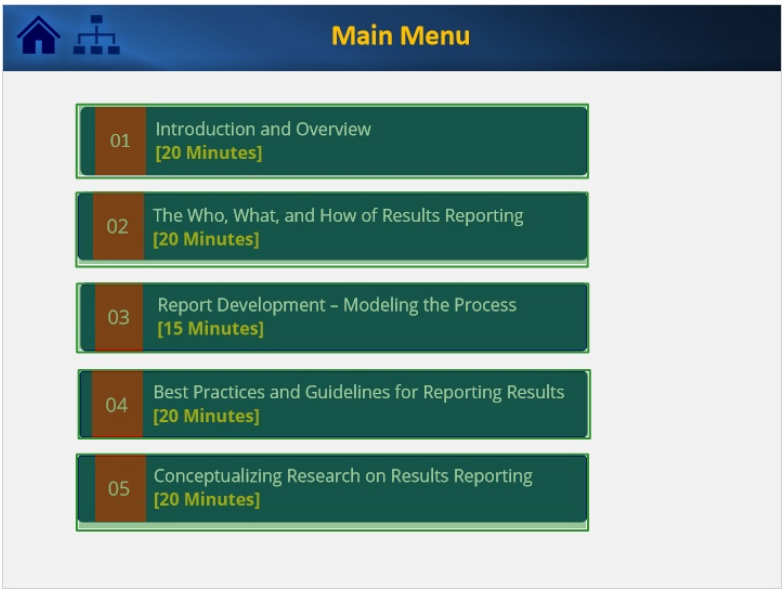
1.11 Prerequisites



1.12 Module Citation & Resource



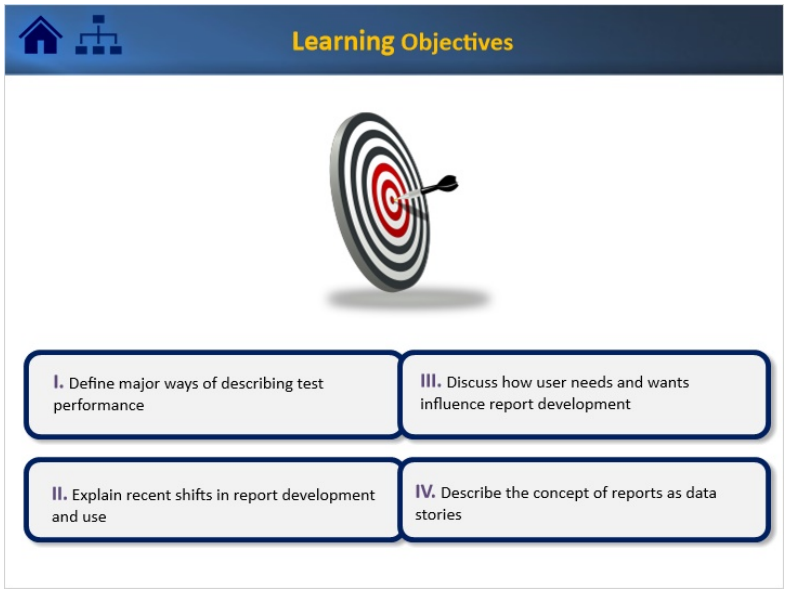
1.13 Main Menu



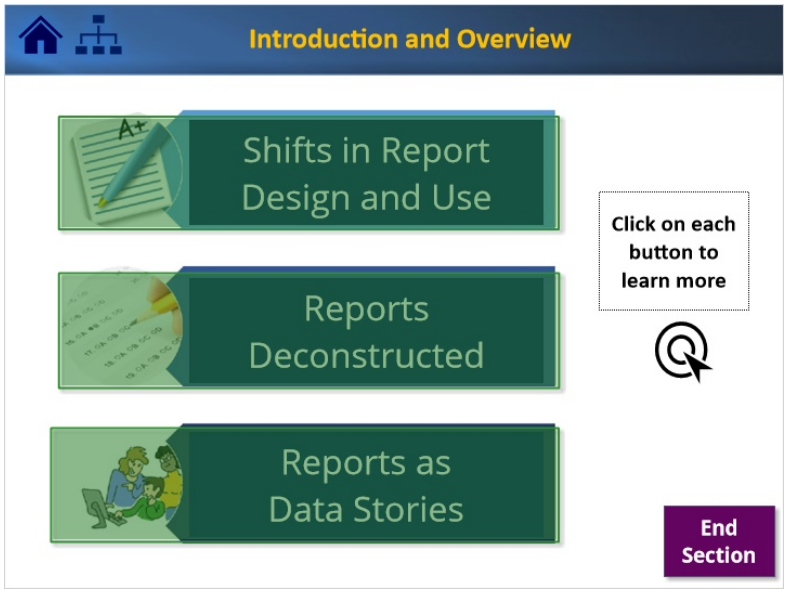
2. Section 1: Introduction & Overview

2.1 Cover: Section 1

2.2 Objectives: Section 1



2.3 Topic Selection

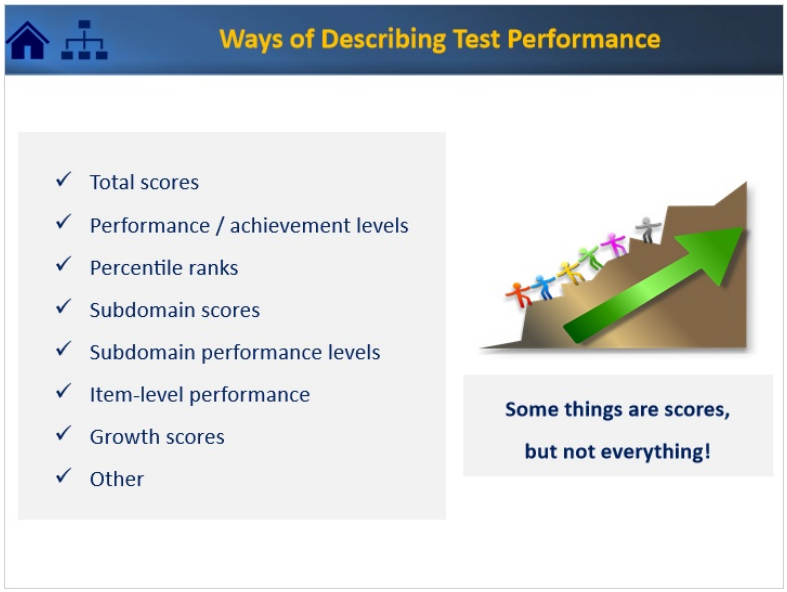


2.4 Bookmark: A Seismic Shift

2.5 A Seismic Shift I

2.6 More about Results Reporting

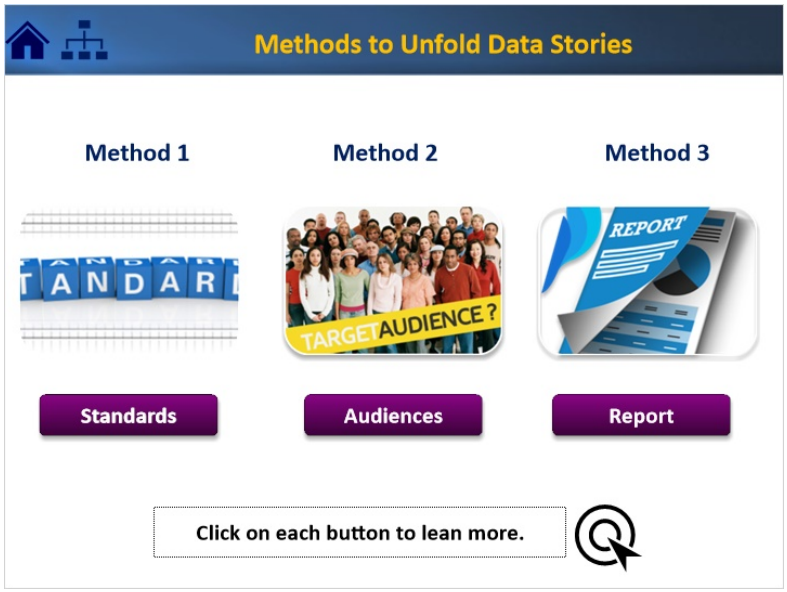
2.7 Ways of Describing Test Performance



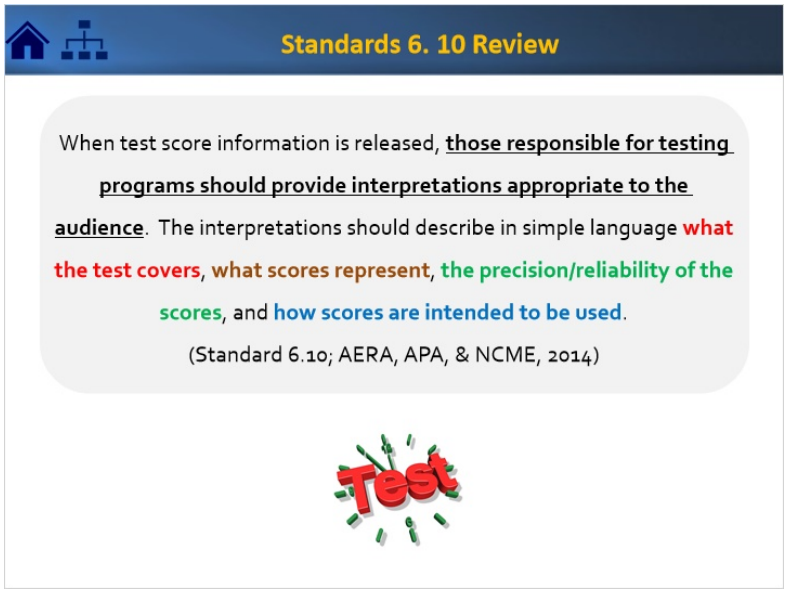
2.8 Score reporting

2.9 A Seismic Shift I

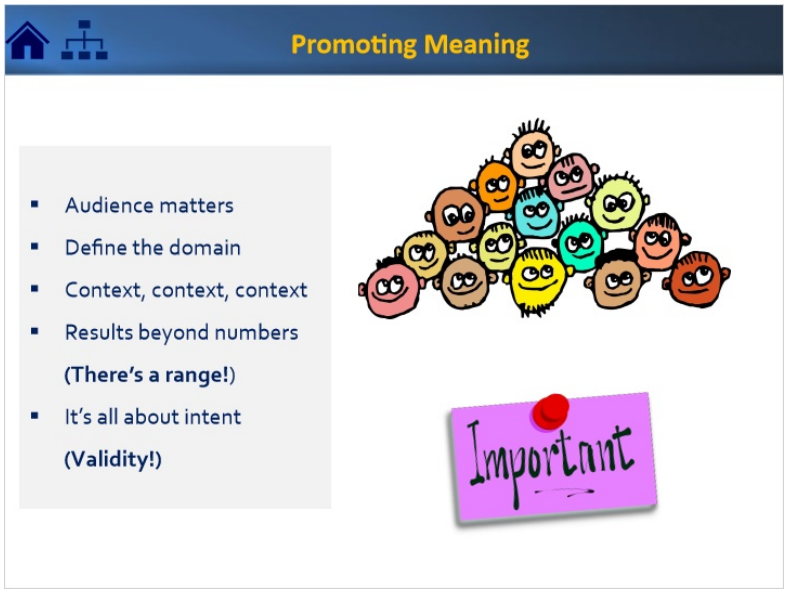
2.10 Topic Selection



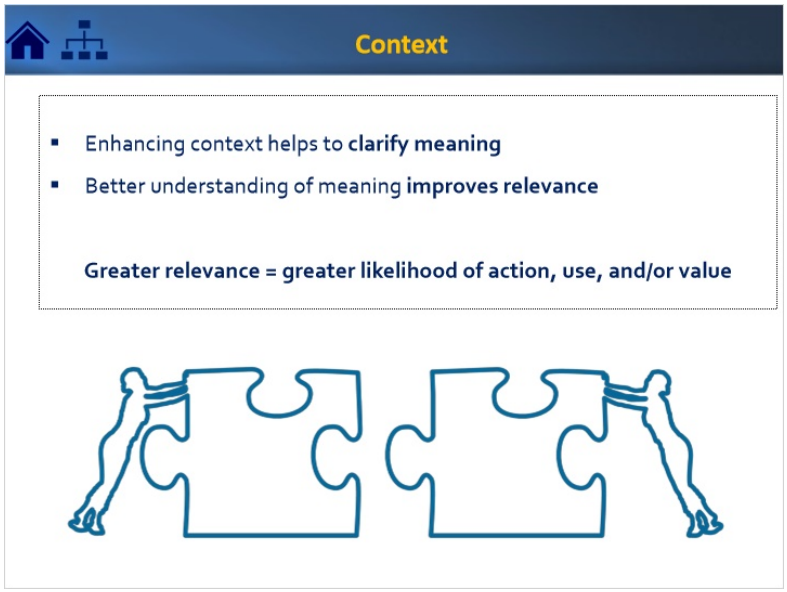
2.11 Standards Overview



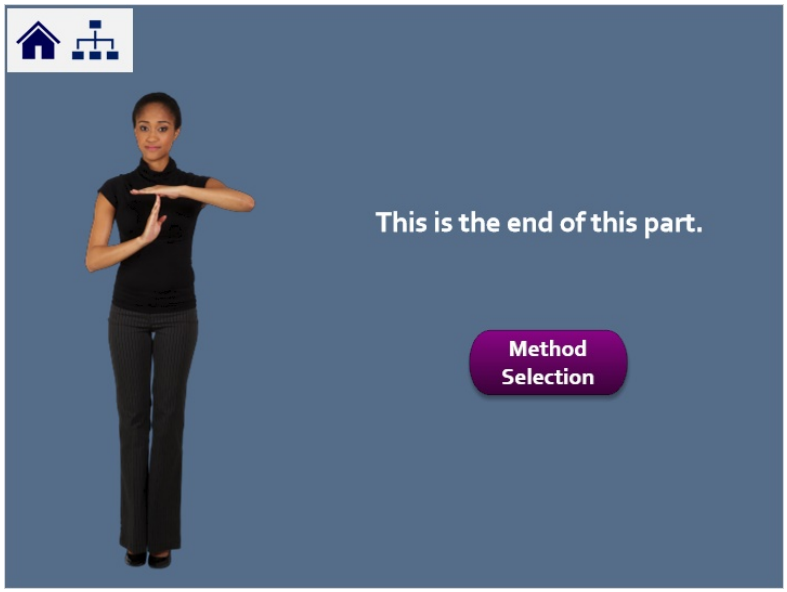
2.12 Promoting Meaning



2.13 Context for Meaning

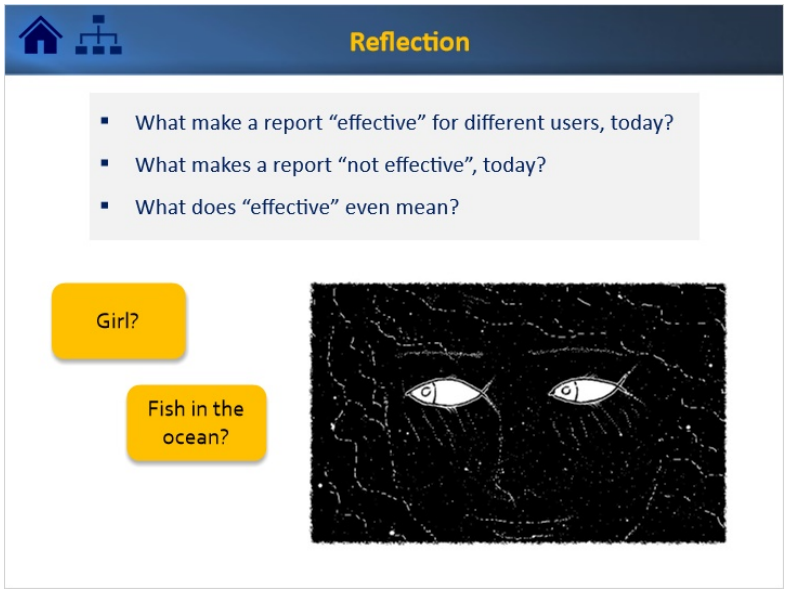


2.14 Bookend: Standard

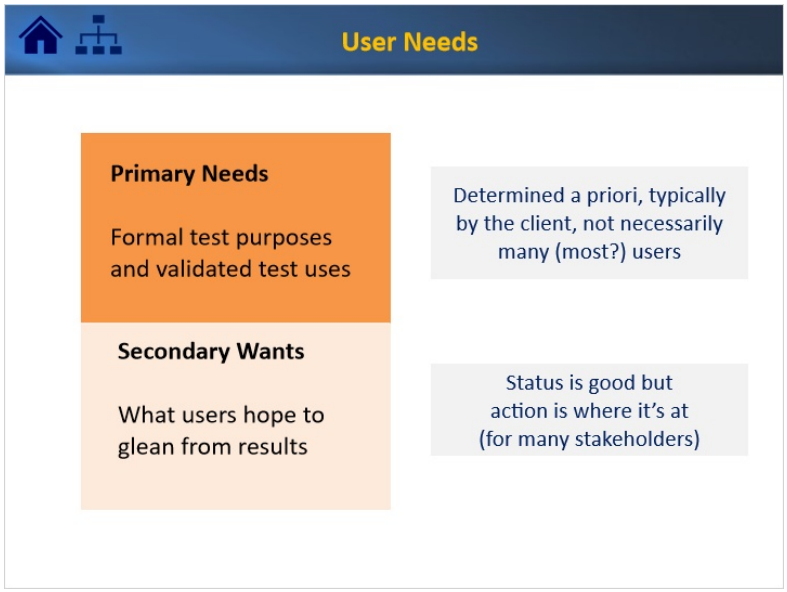


2.15 Report Overview

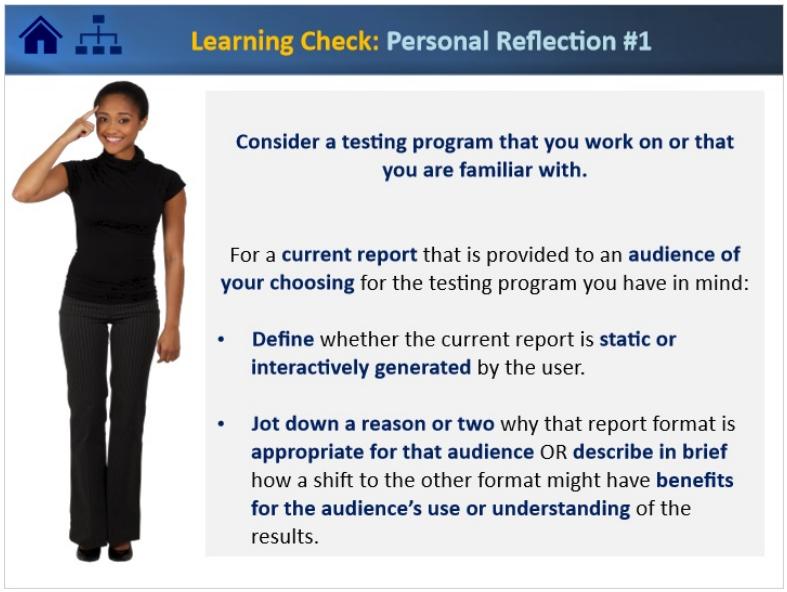
2.16 Reflection



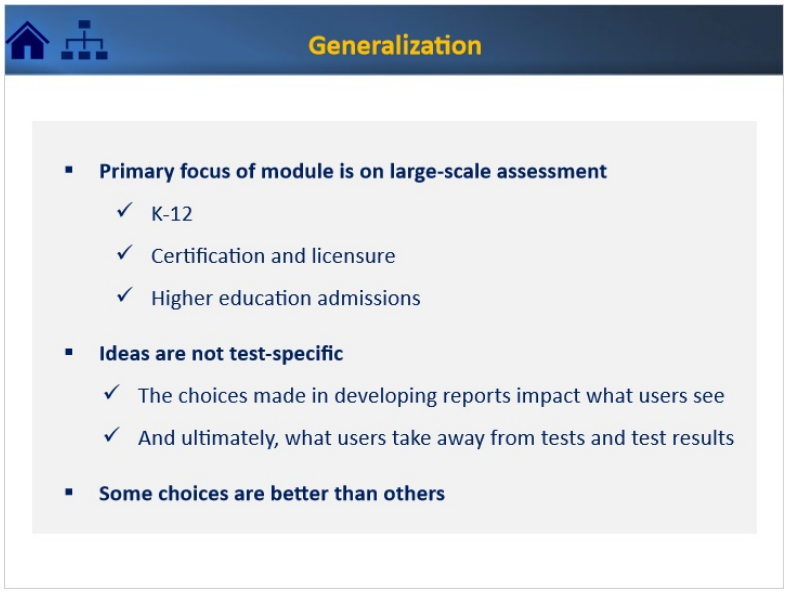
2.17 User needs



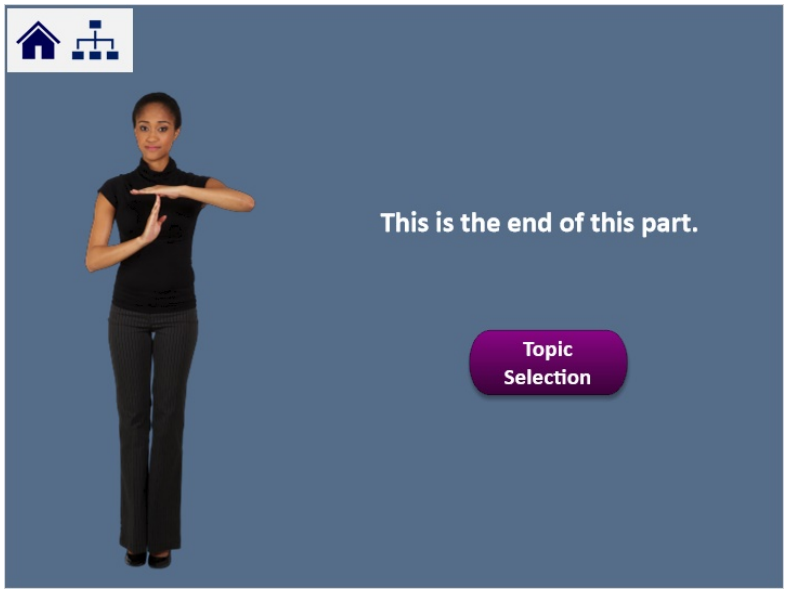
2.18 Personal Reflection #1



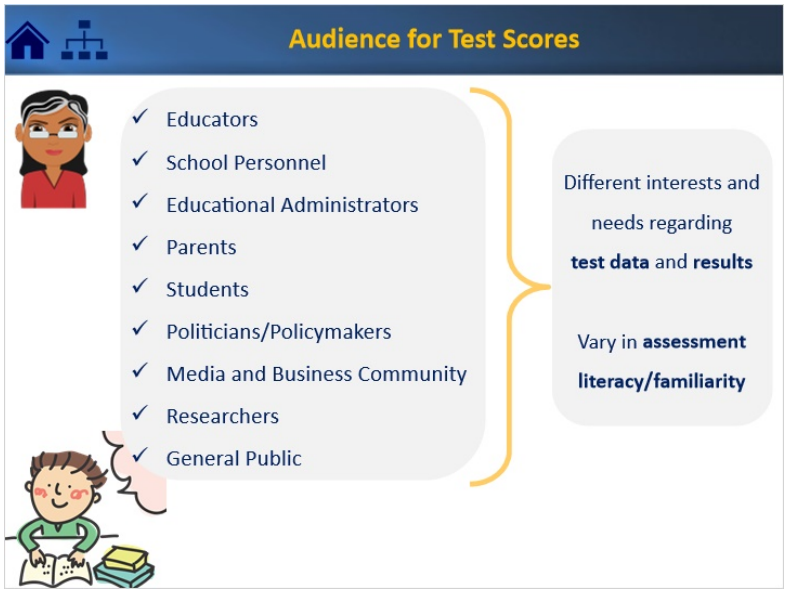
2.19 Boundaries



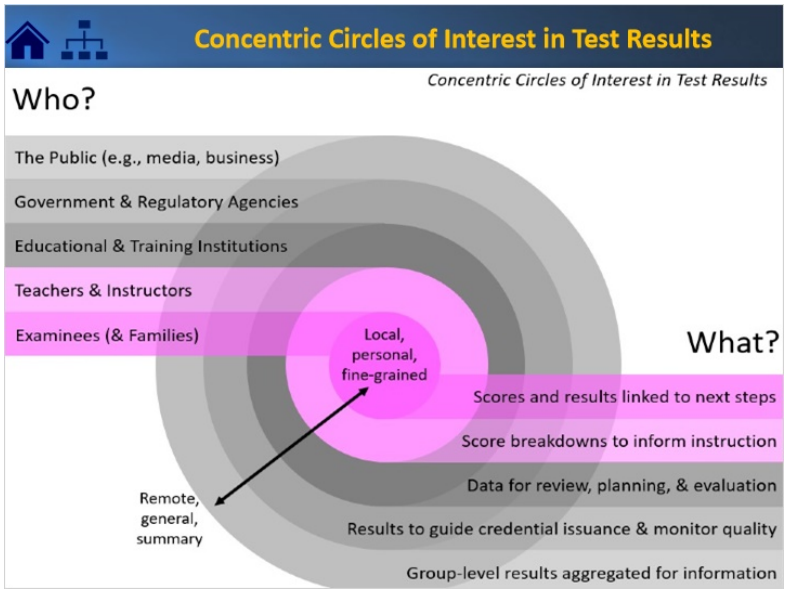
2.20 Bookend: Report



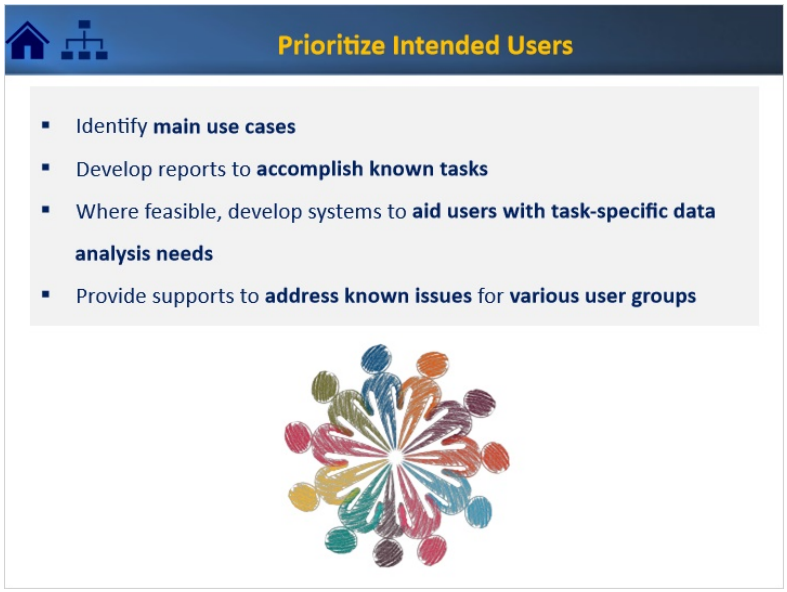
2.21 Audience Overview



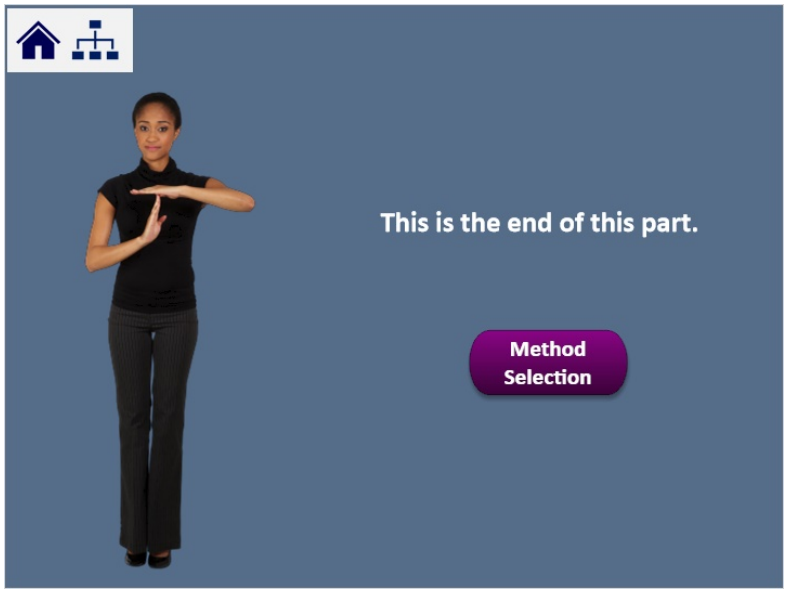
2.22 Circles of Interest



2.23 Priortize Intended Users



2.24 Bookend: Audience

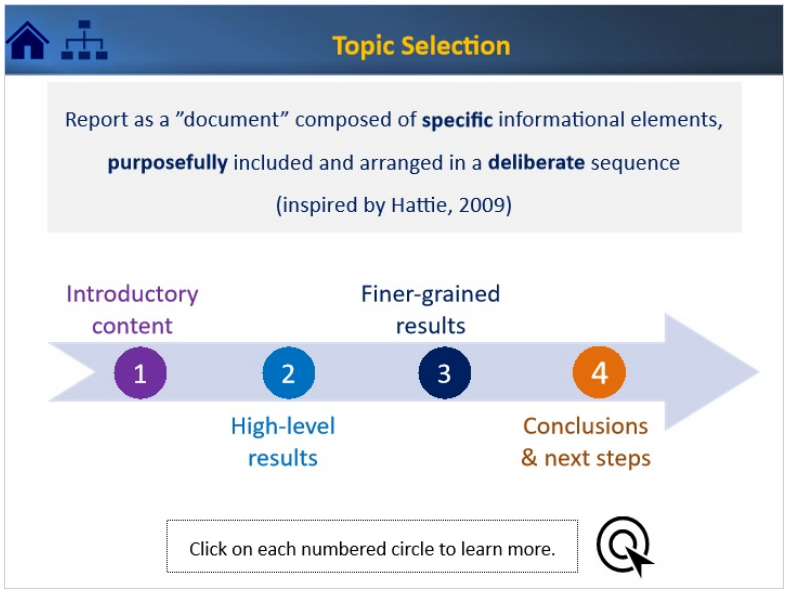


2.25 Bookmark: Deconstructing a Report

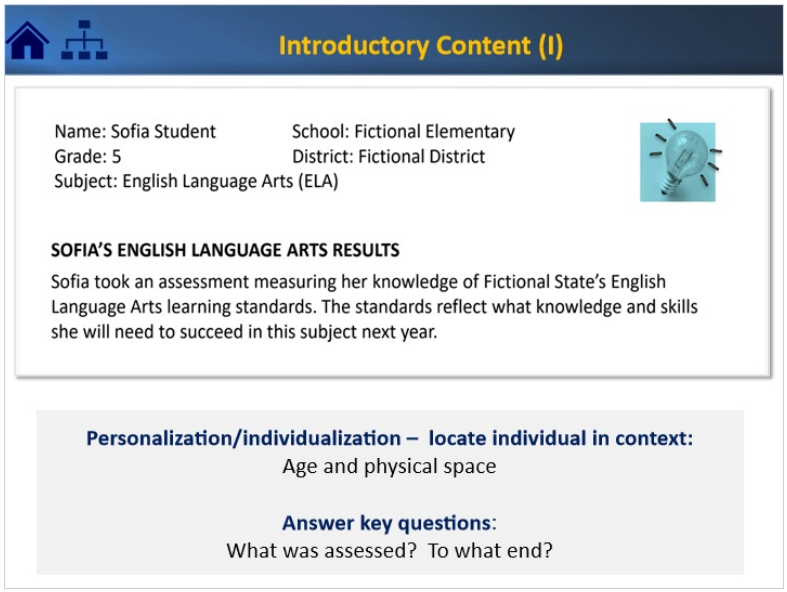
2.26 What is a Report?

2.27 Deconstructing a report II

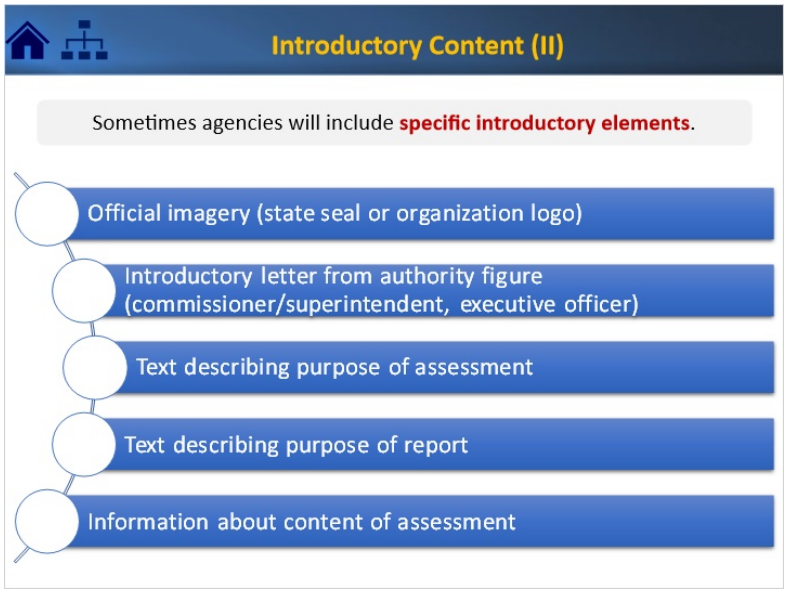
2.28 Topic Selection



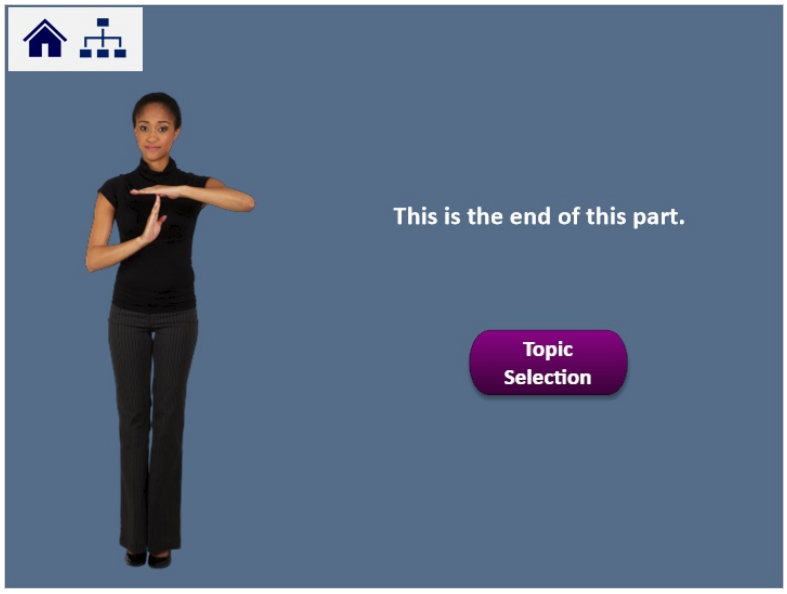
2.29 Introductory Content (I)



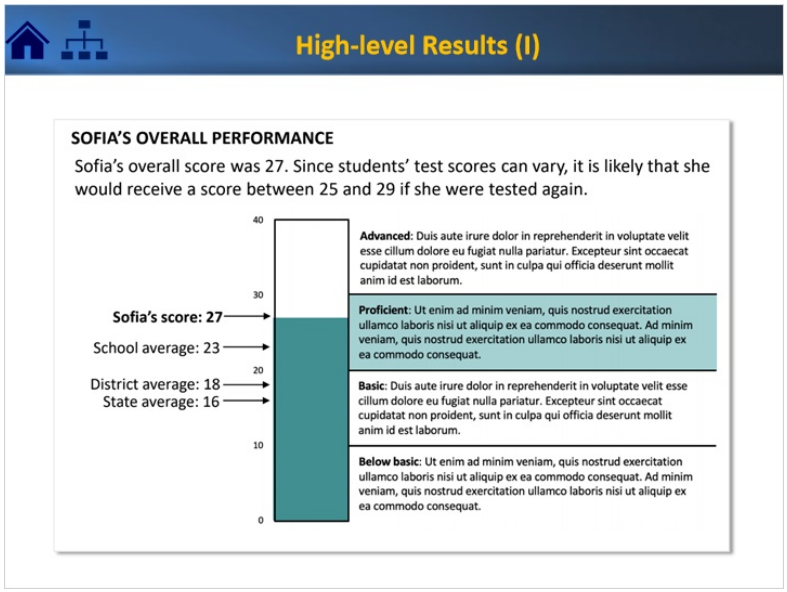
2.30 Introductory Content (II)



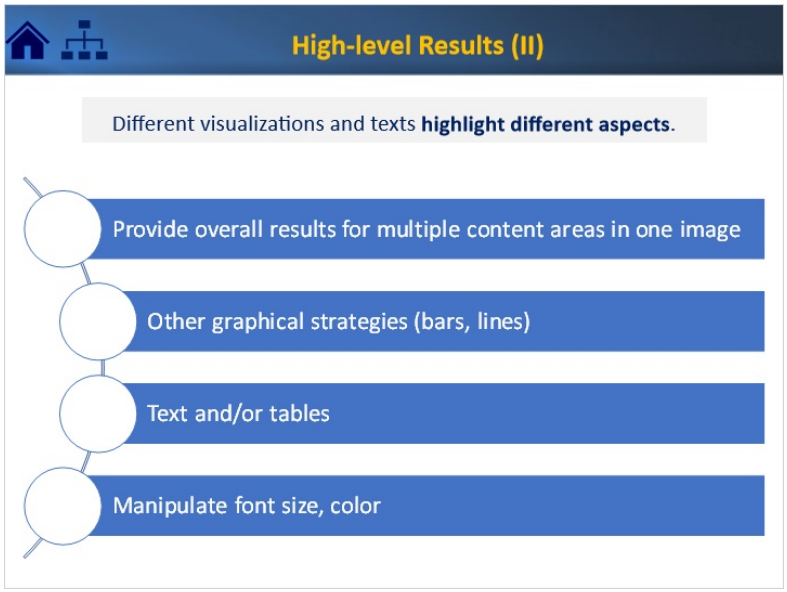
2.31 Bookend: Introductory Content



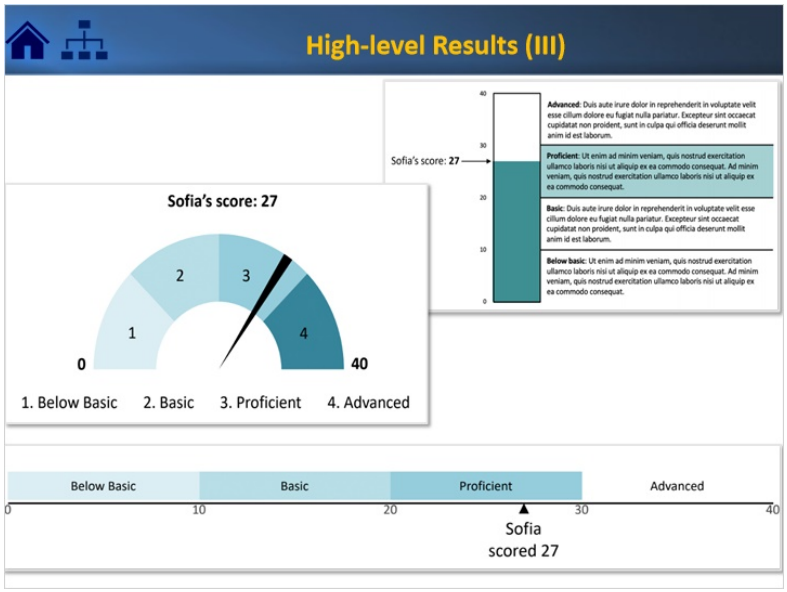
2.32 High-level Results (I)



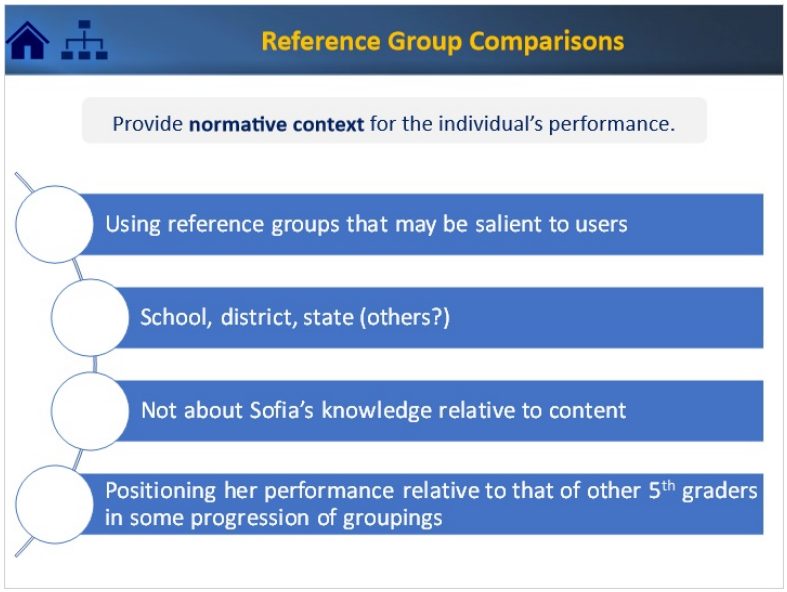
2.33 High-level Results (II)



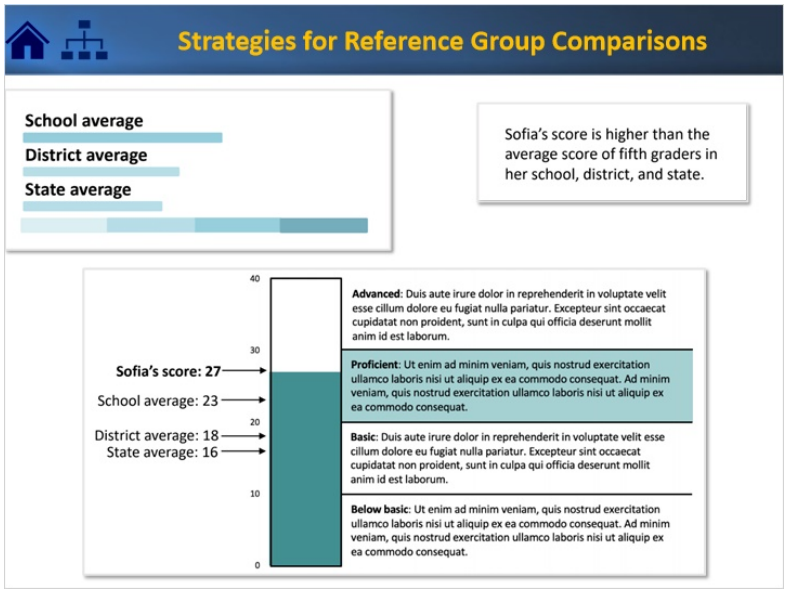
2.34 High-level Results (III)



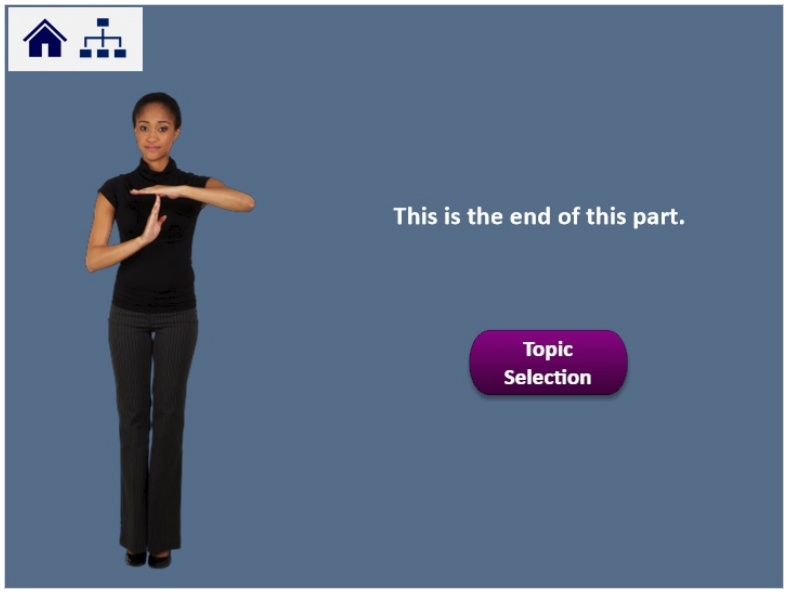
2.35 High-level Results (IV)



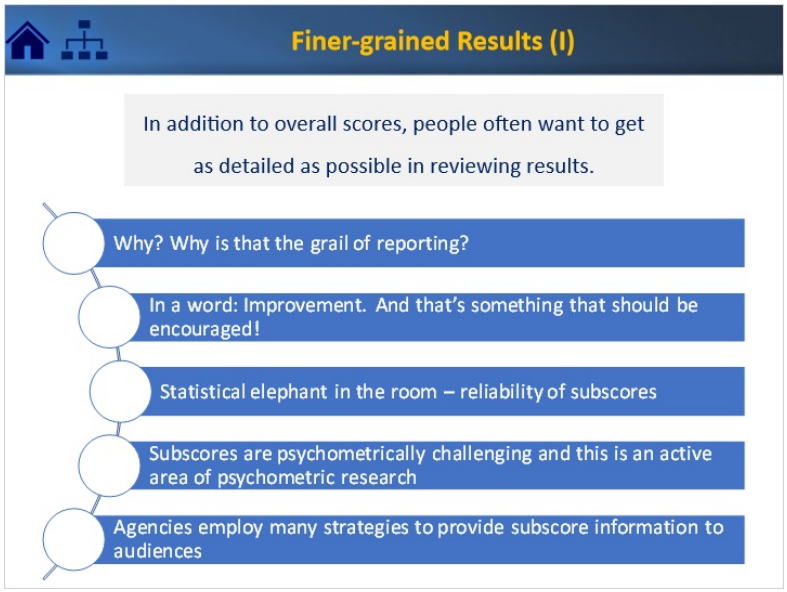
2.36 High-level Results (V)



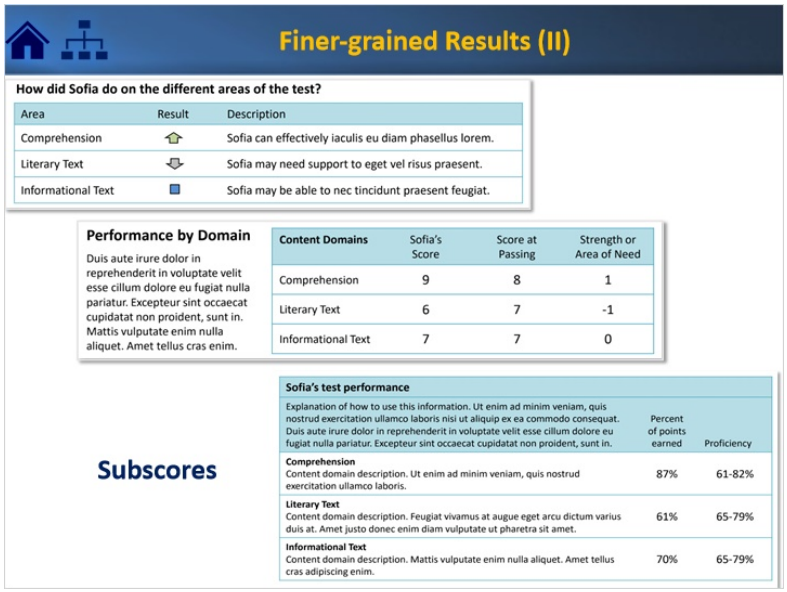
2.37 Bookend: Higher-level Results



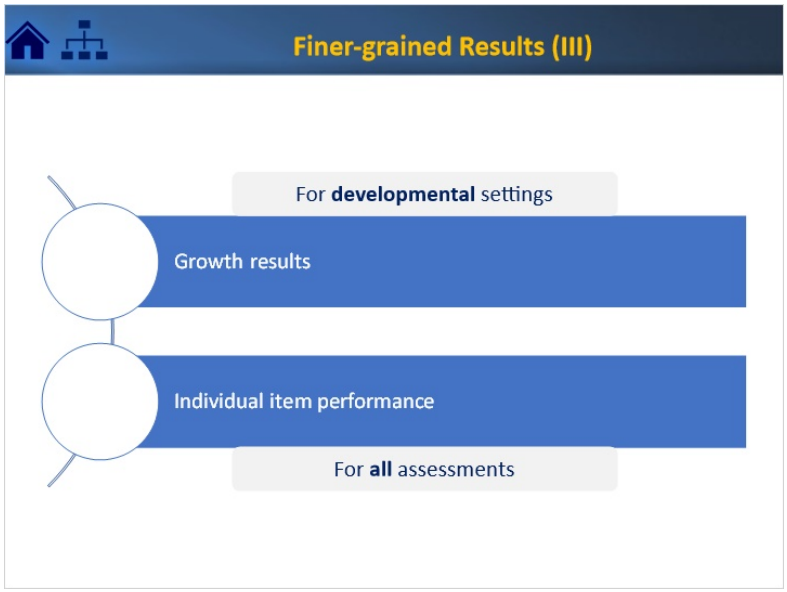
2.38 Finer-grained Results (I)



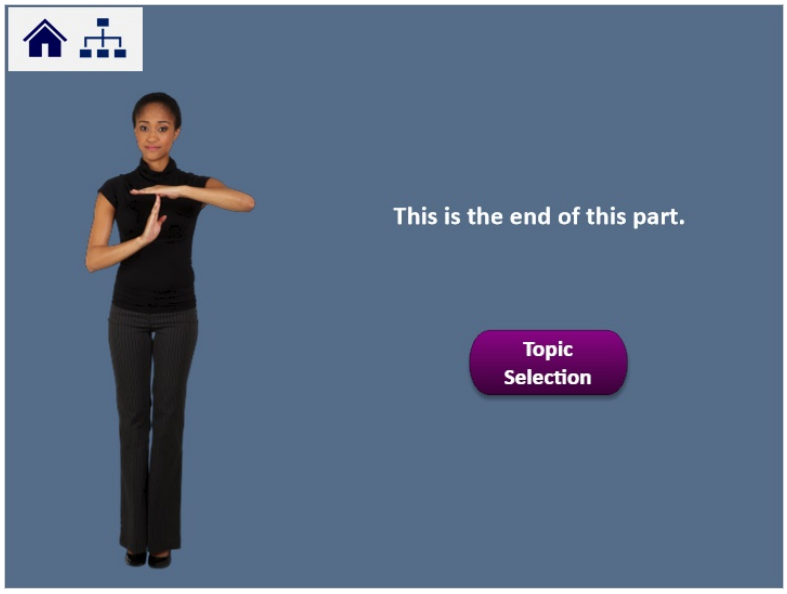
2.39 Finer-grained Results (II)



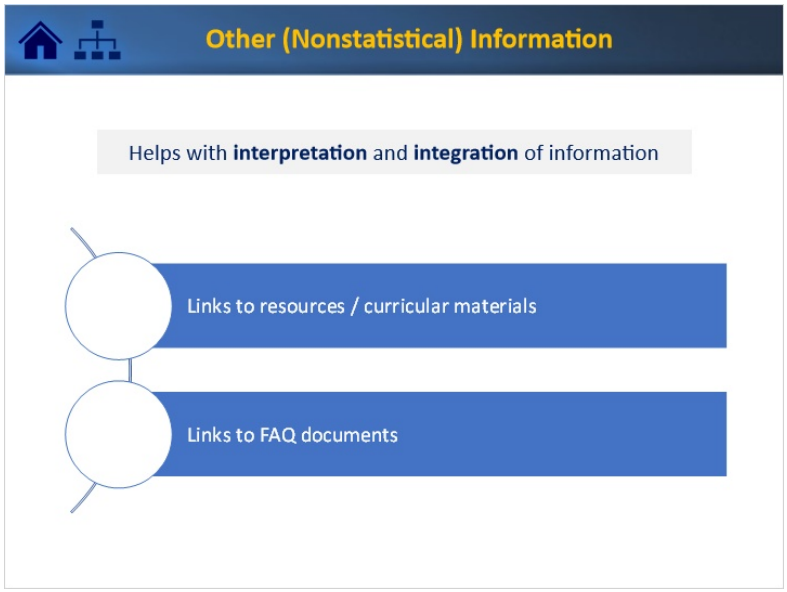
2.40 Finer-grained Results (III)



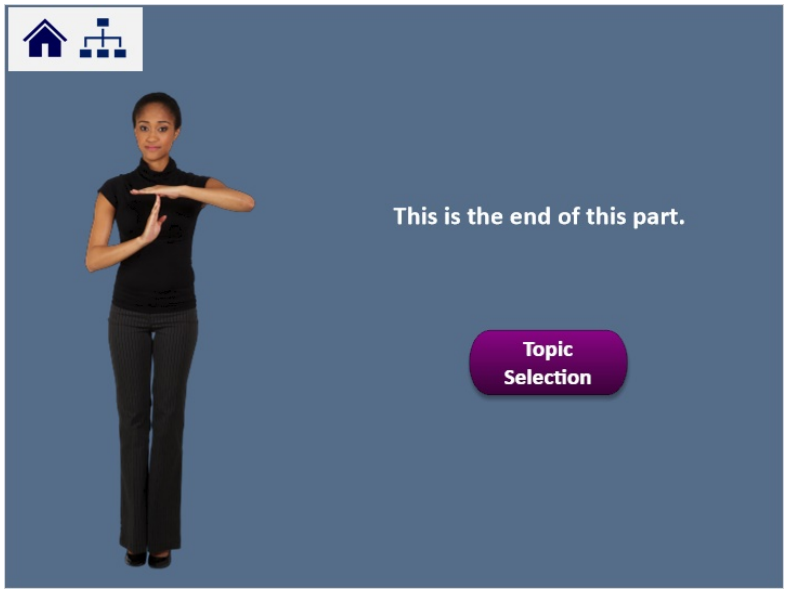
2.41 Bookend: Finer-grained Results



2.42 Other Information

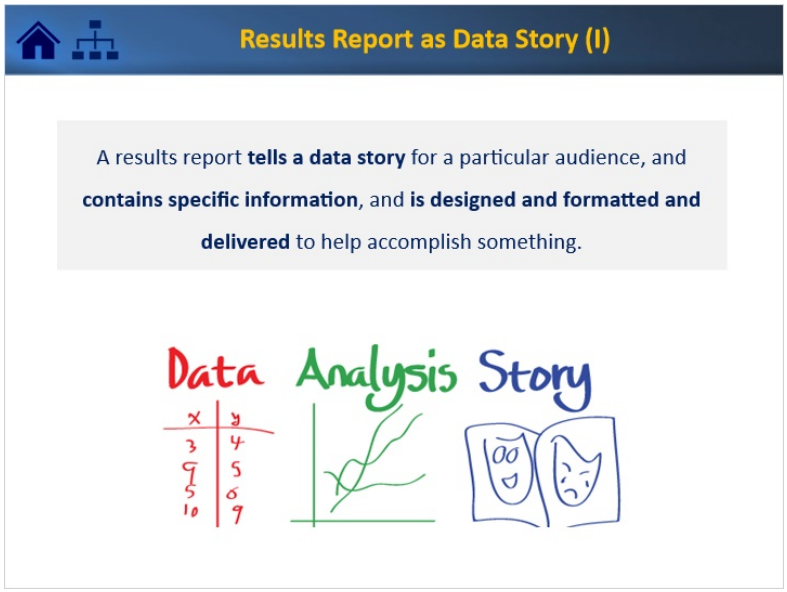


2.43 Bookmark: Other Information

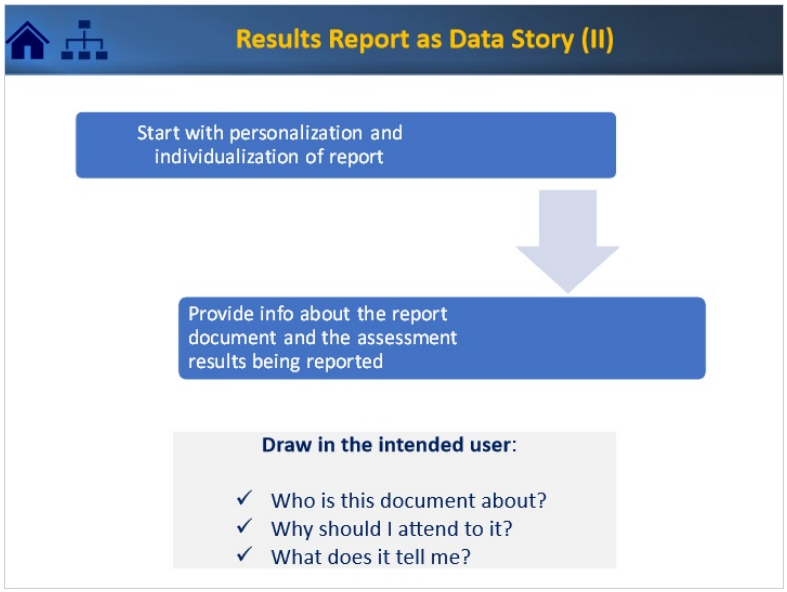


2.44 Bookmark: Reports as Data Stories

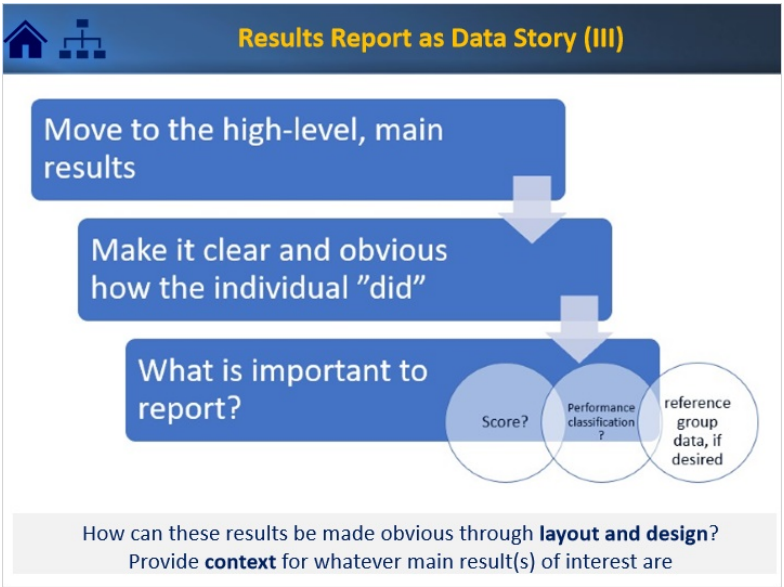
2.45 Results Report as Data Story (I)



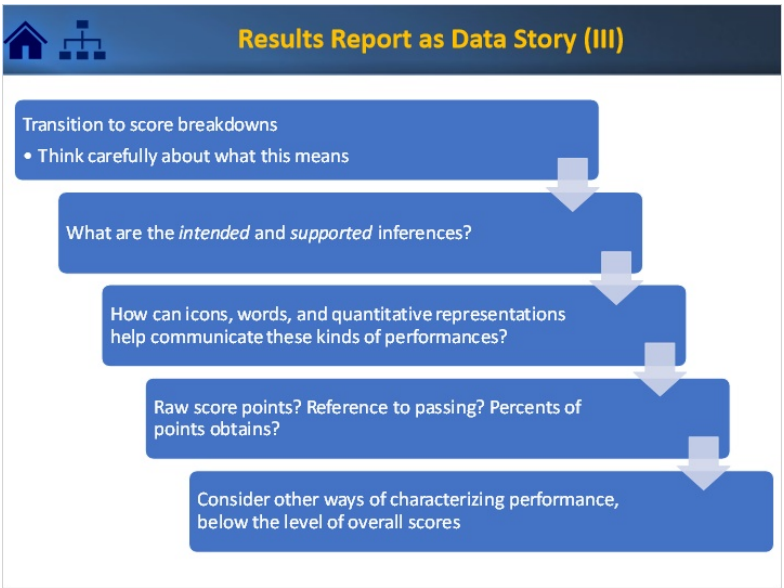
2.46 Results Report as Data Story (II)



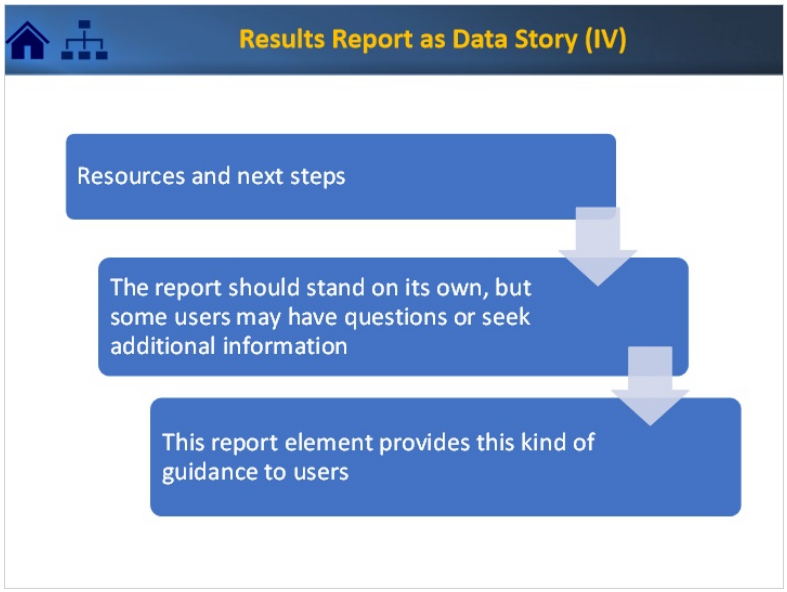
2.47 Results Report as Data Story (III)



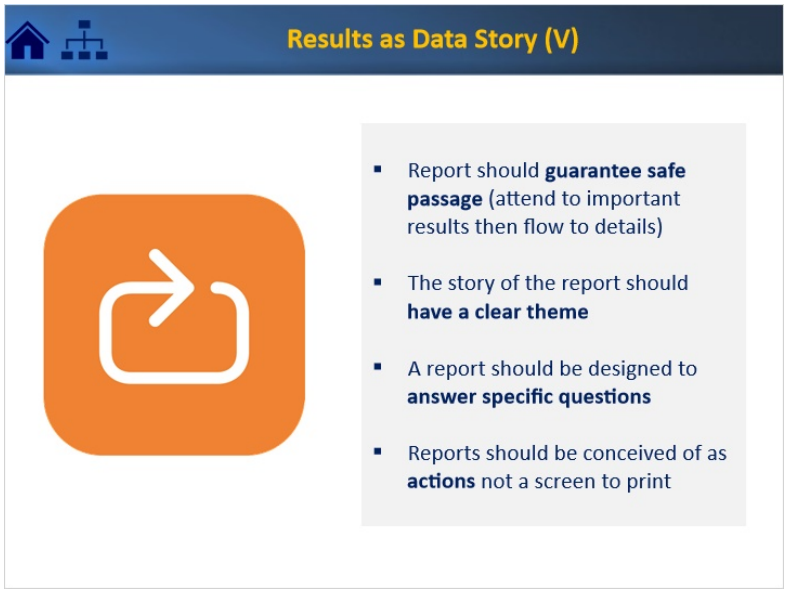
2.48 Results Report as Data Story (III)



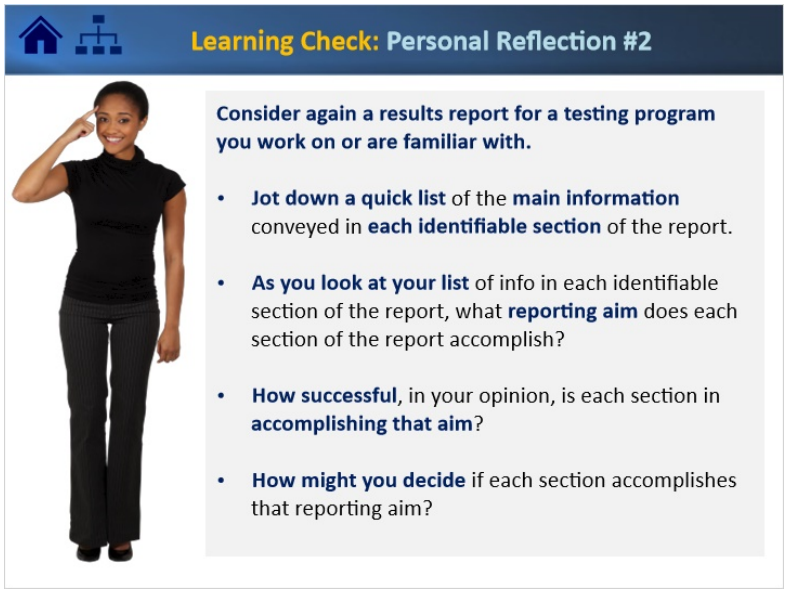
2.49 Results Report as Data Story (IV)



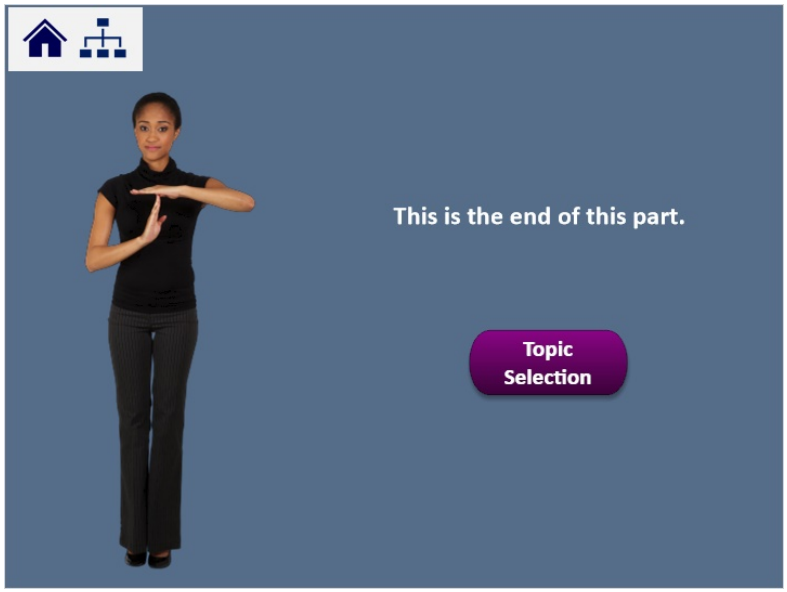
2.50 Results as Data Story (V)



2.51 Personal Reflection #2

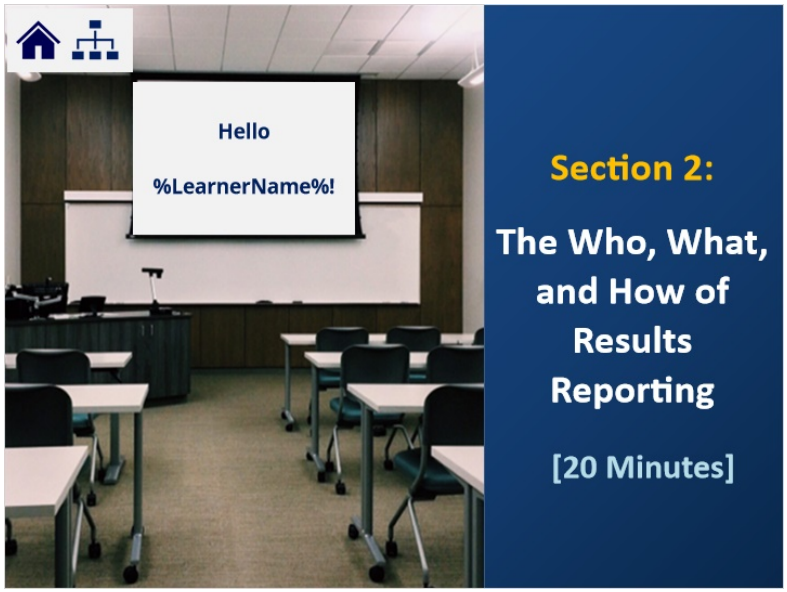


2.52 End of Section 1

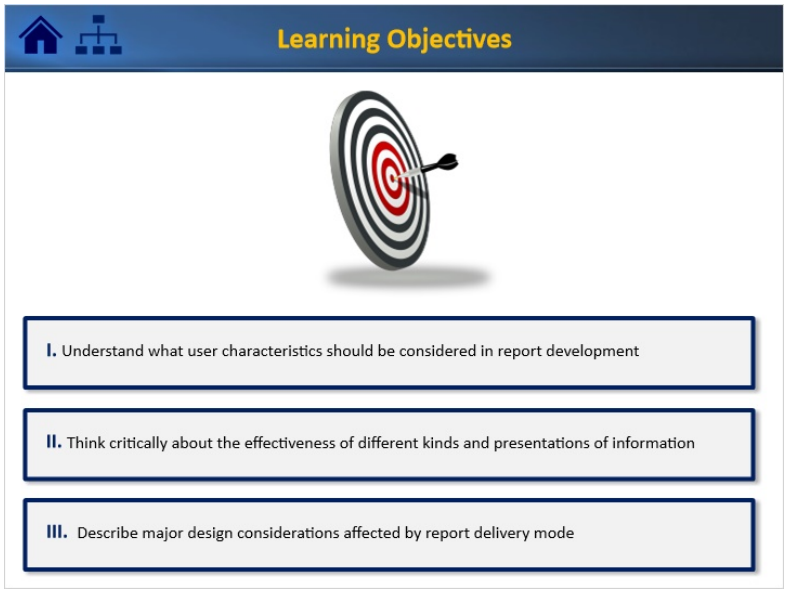


3. Section 2: The Who, What, and How of Results Reporting

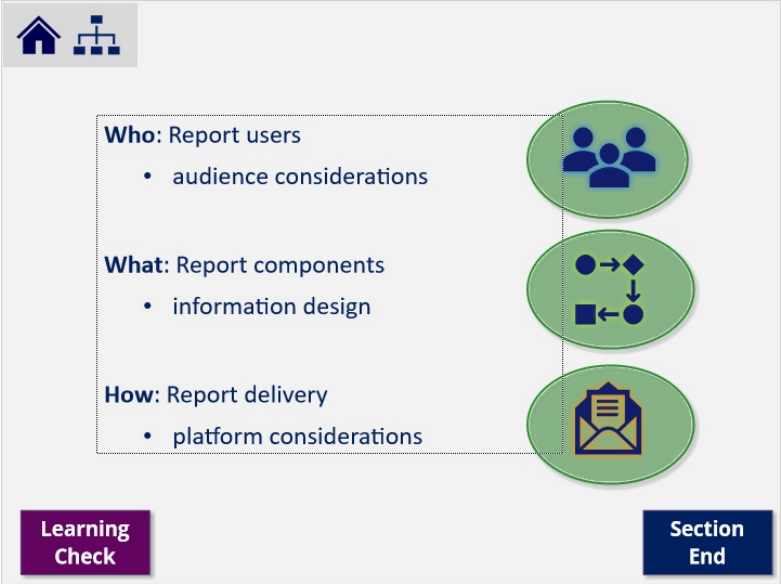
3.1 Cover: Section 2



3.2 Objectives: Section 2

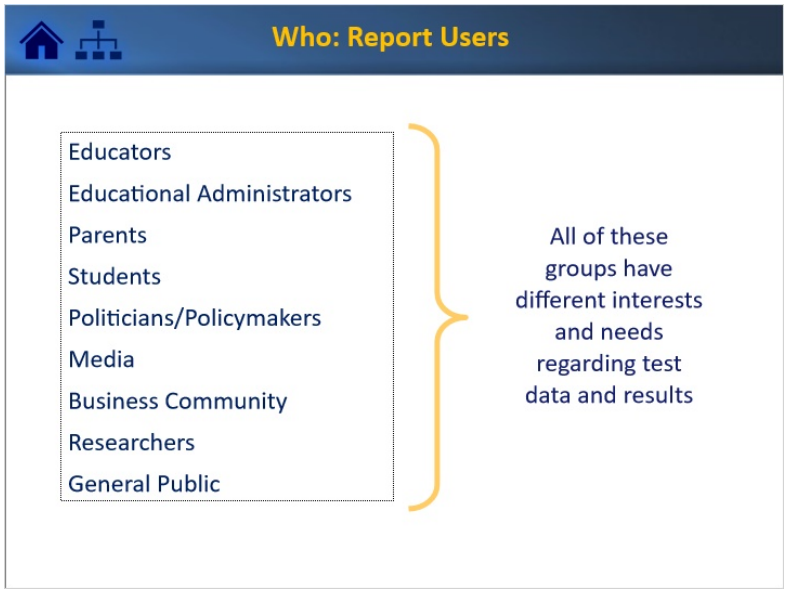


3.3 Topic Selection

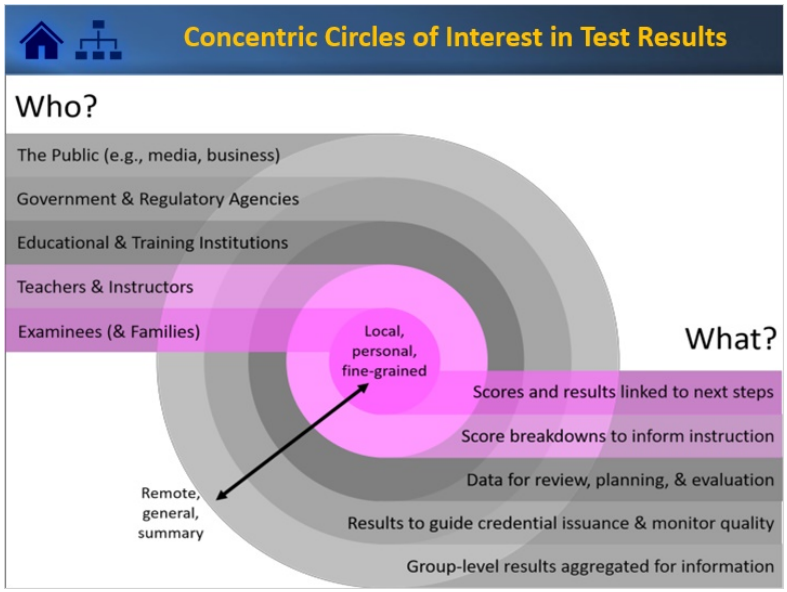


3.4 Bookmark:Who

3.5 Who I



3.6 Who II



3.7 Who III

Considering User Characteristics

Zapata-Rivera and Katz (2014) recommend conducting an **audience analysis** to connect the score report design process and report specifics to audience characteristics, noting:

"Although most design guidelines focus on making score reports understandable to people who are not testing professionals, audiences should be defined by more than just their lack of statistical knowledge.

3.8 Who IV

Audience Analysis (Zapata-Rivera & Katz, 2014)

Needs

- What are the users' goals in using the report?
- What do they want to accomplish?

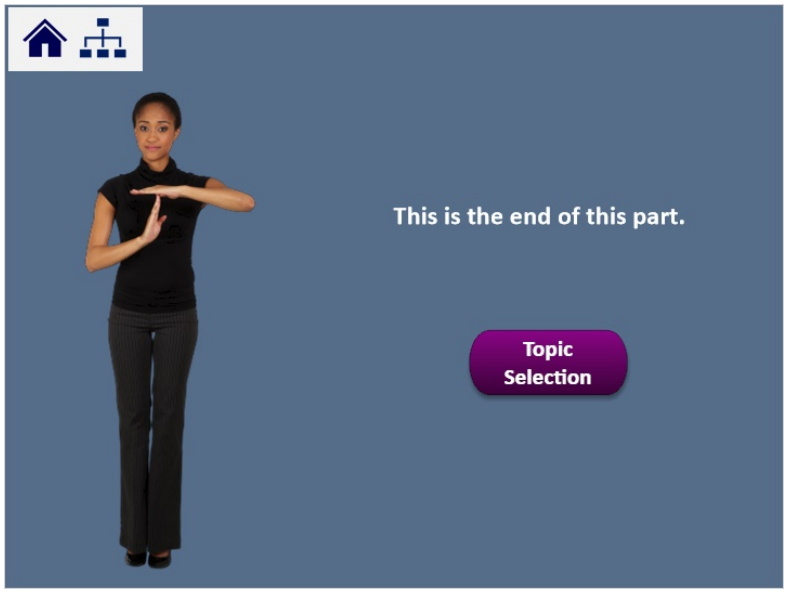
Knowledge

- What do users already know about the assessment and its context?
- What knowledge gaps might impact their understanding?

Attitudes

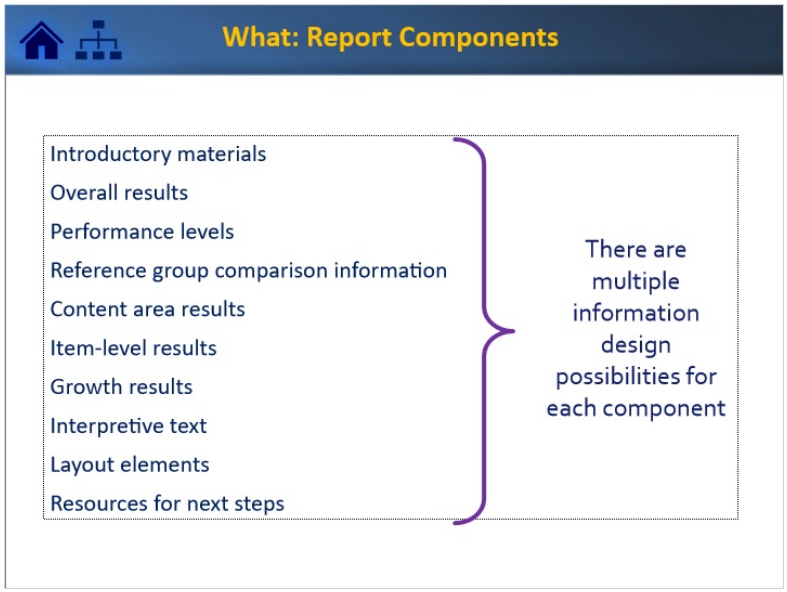
- What are users' perceptions of the assessment?
- How might the design of the report contribute to a productive attitude about the results?

3.9 Bookend: Who

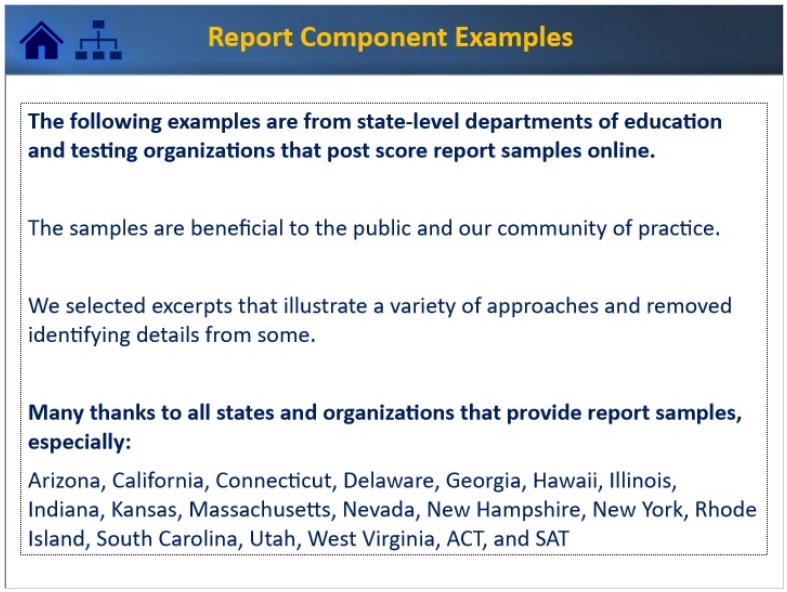


3.10 Bookmark: What

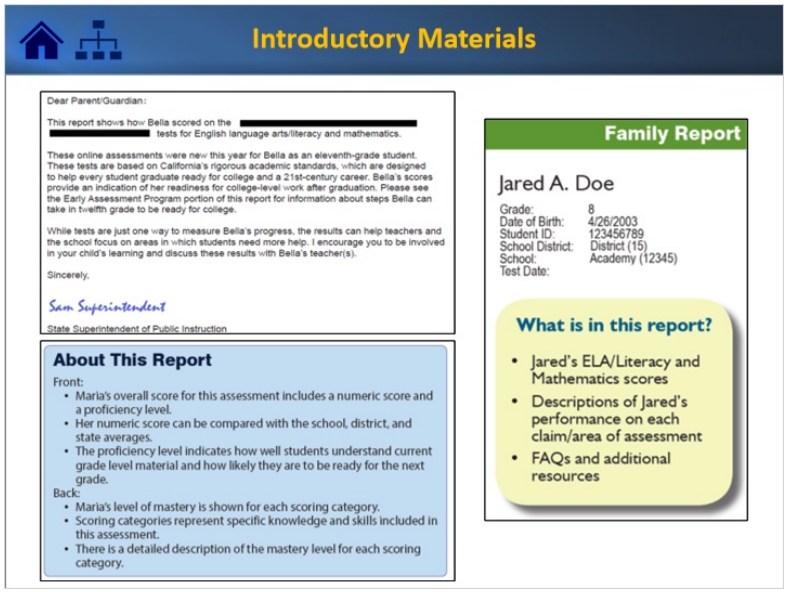
3.11 What I



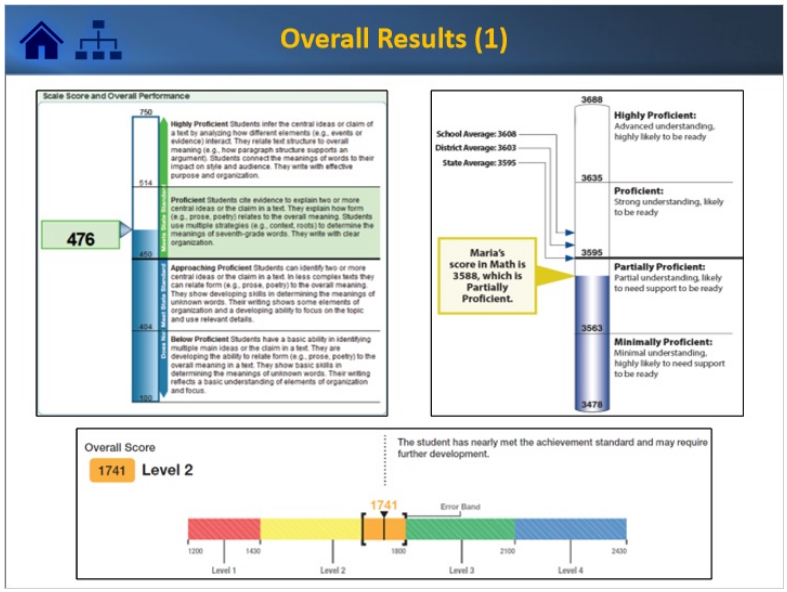
3.12 What II



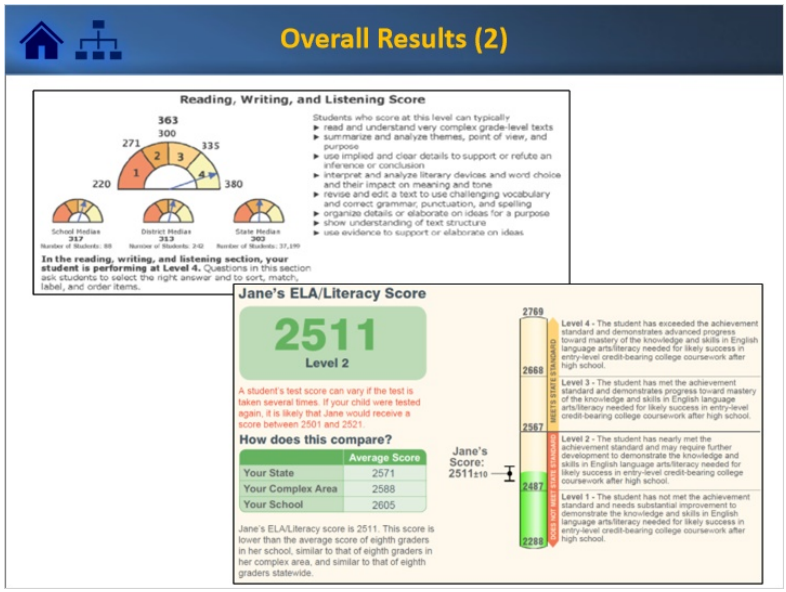
3.13 What III



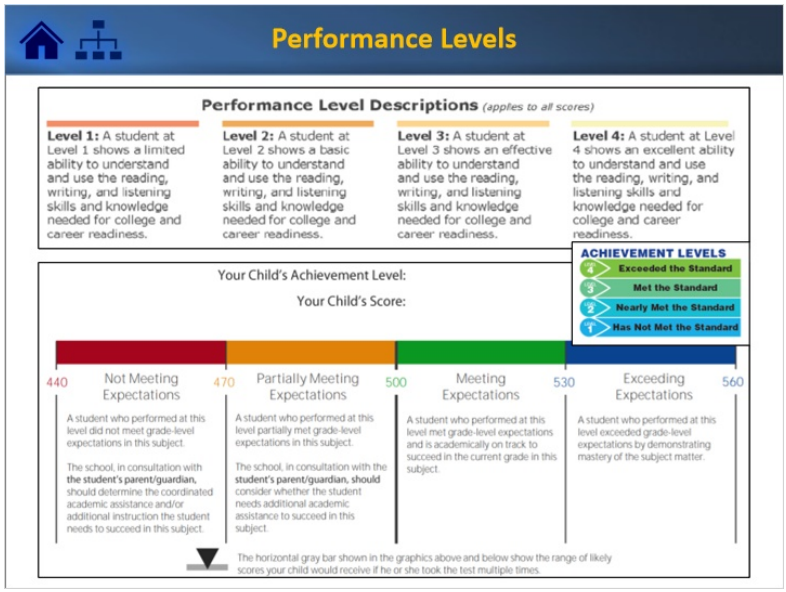
3.14 What IV



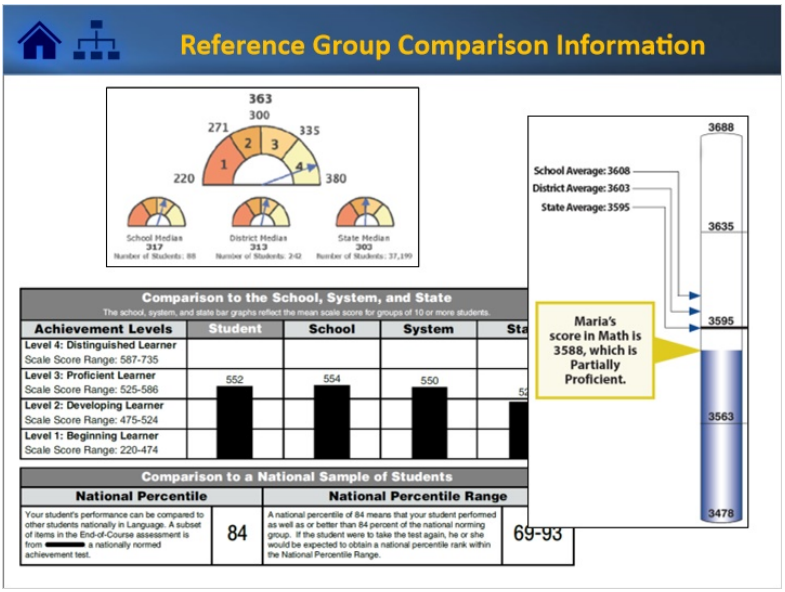
3.15 What V



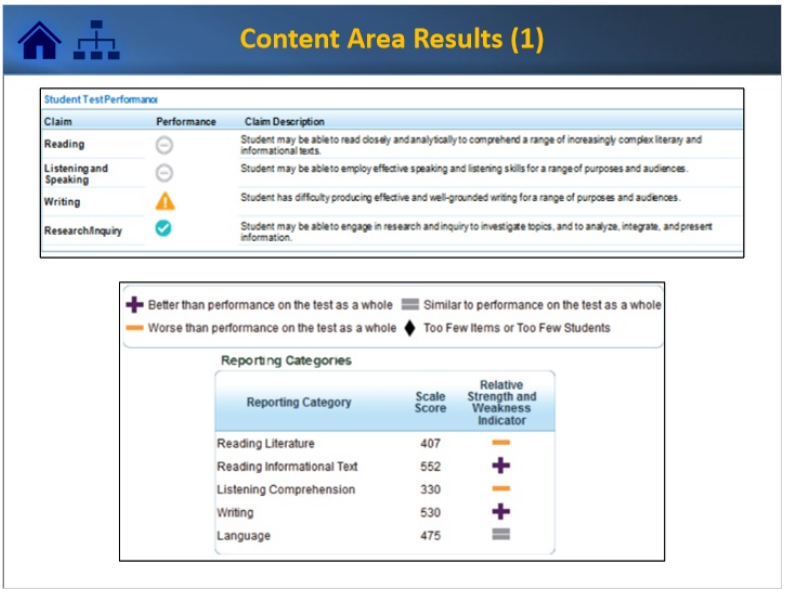
3.16 What VI



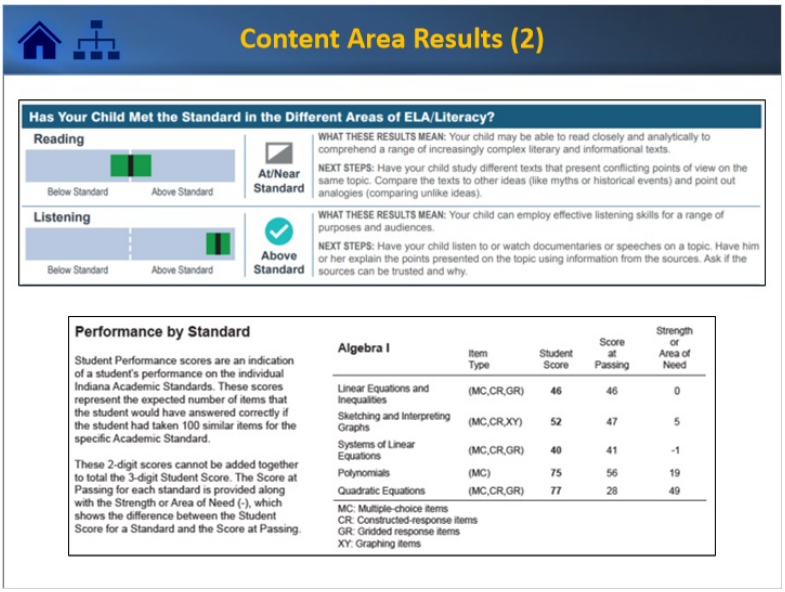
3.17 What VII



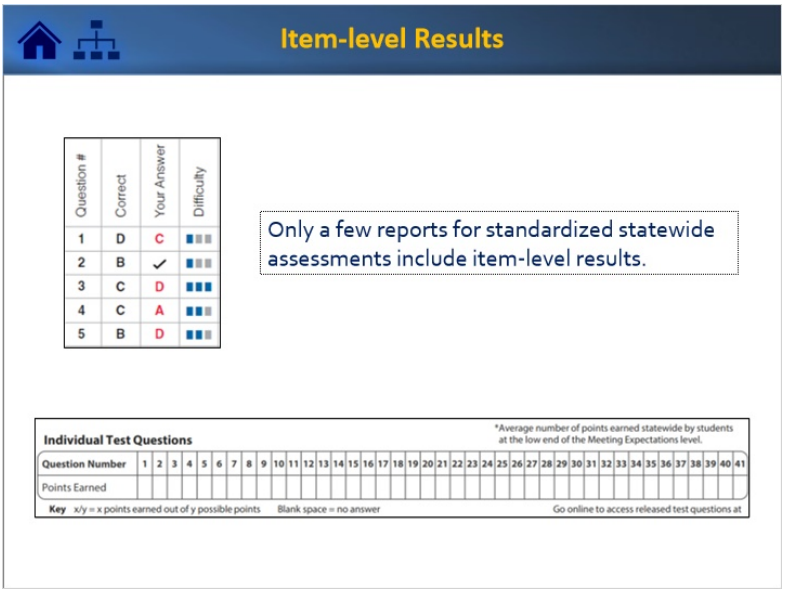
3.18 What VIII



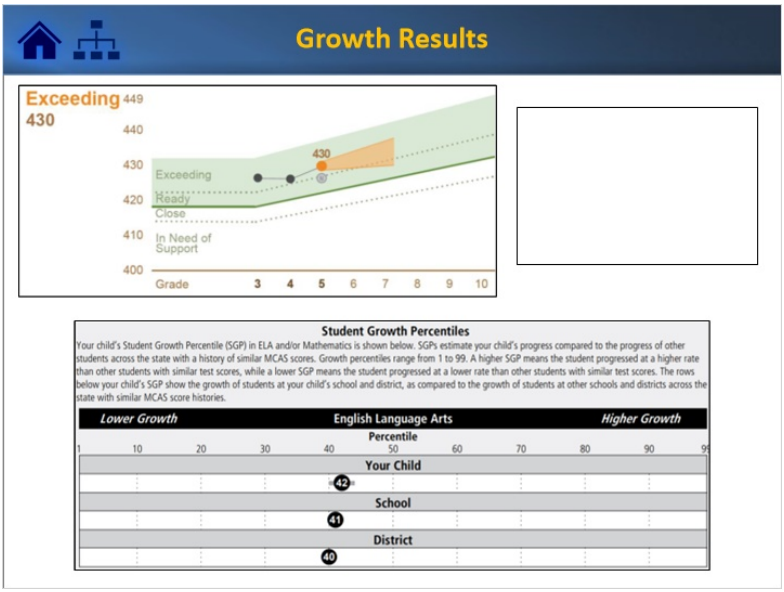
3.19 What IX



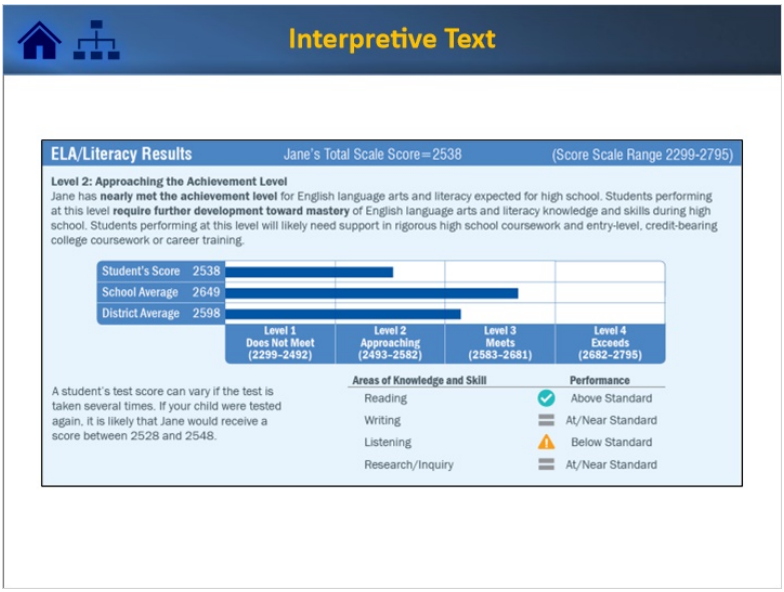
3.20 What X



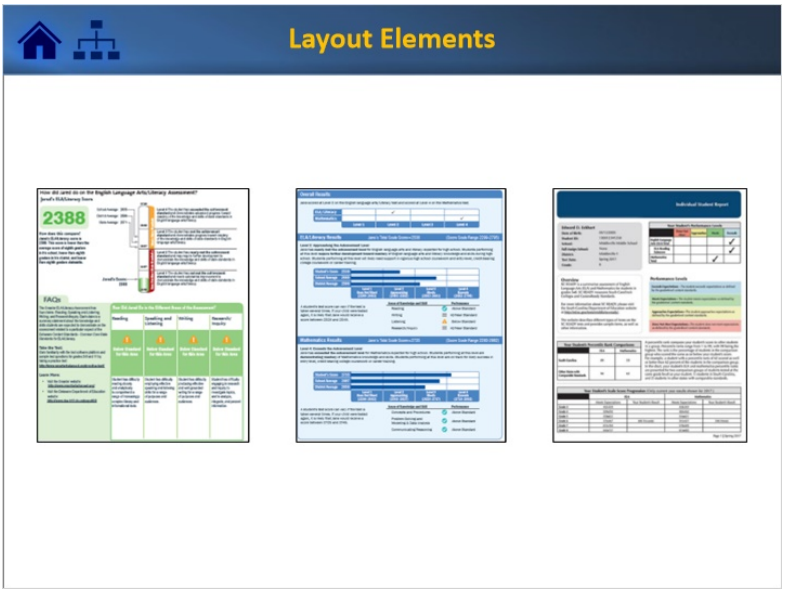
3.21 What XI



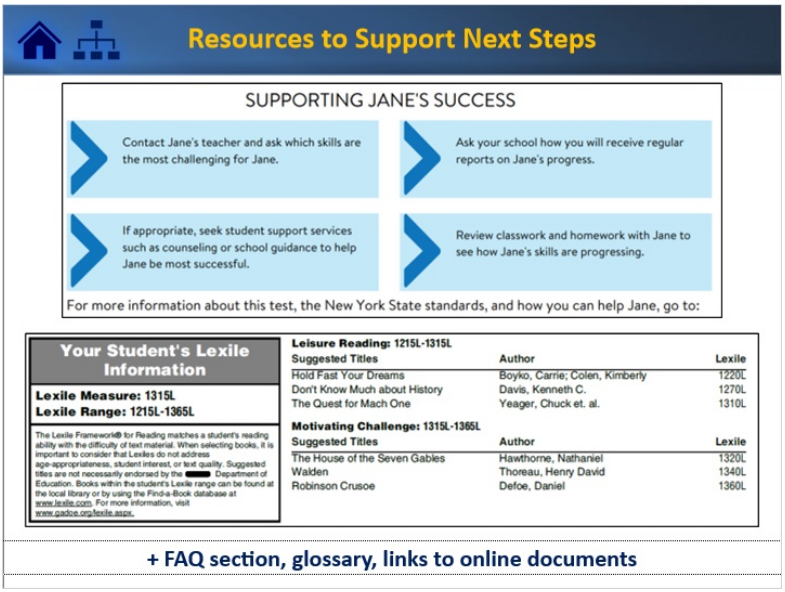
3.22 What XII



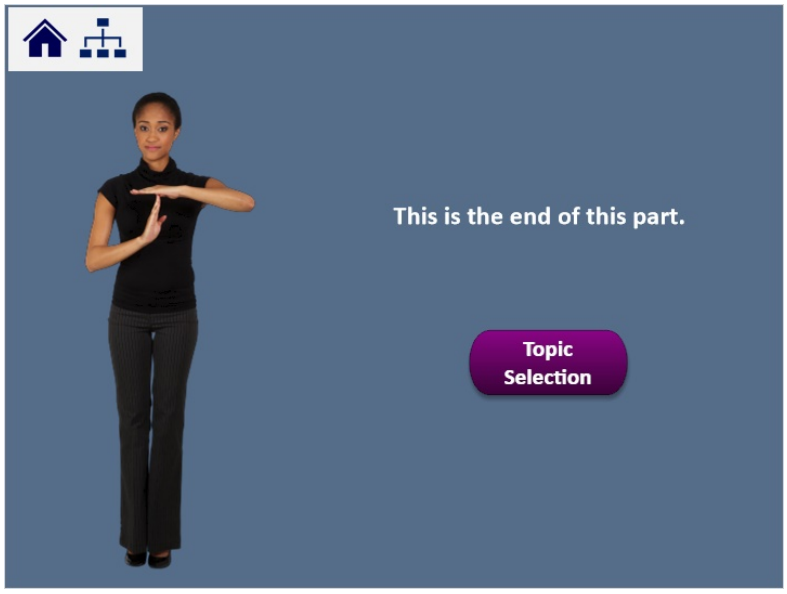
3.23 What XIII



3.24 What XIV

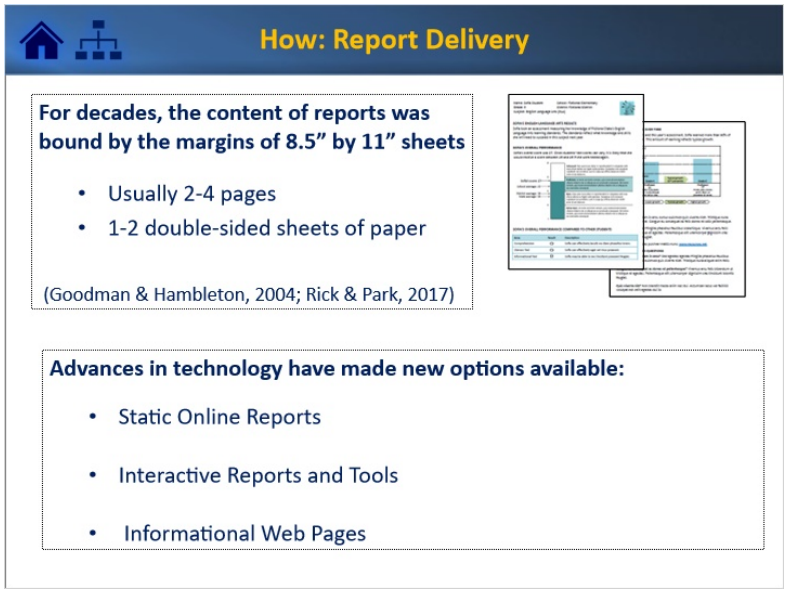


3.25 Bookend: What

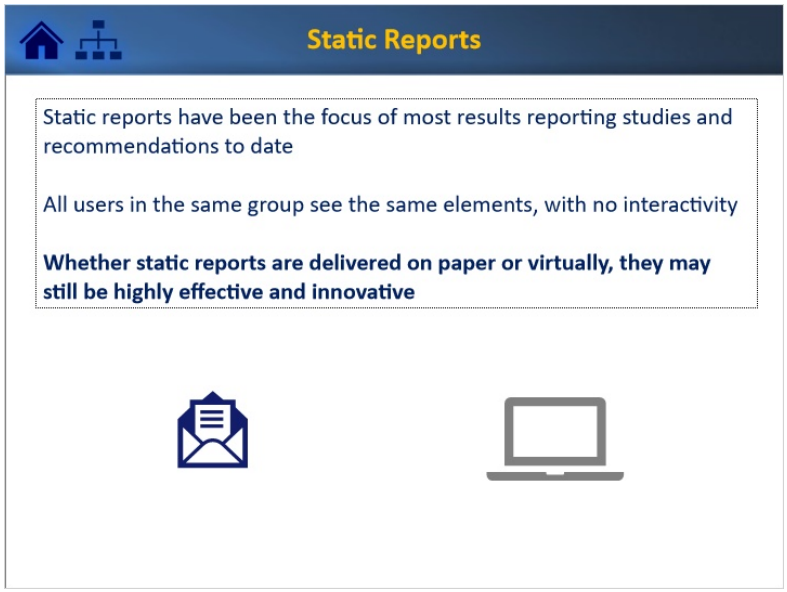


3.26 Bookmark: How

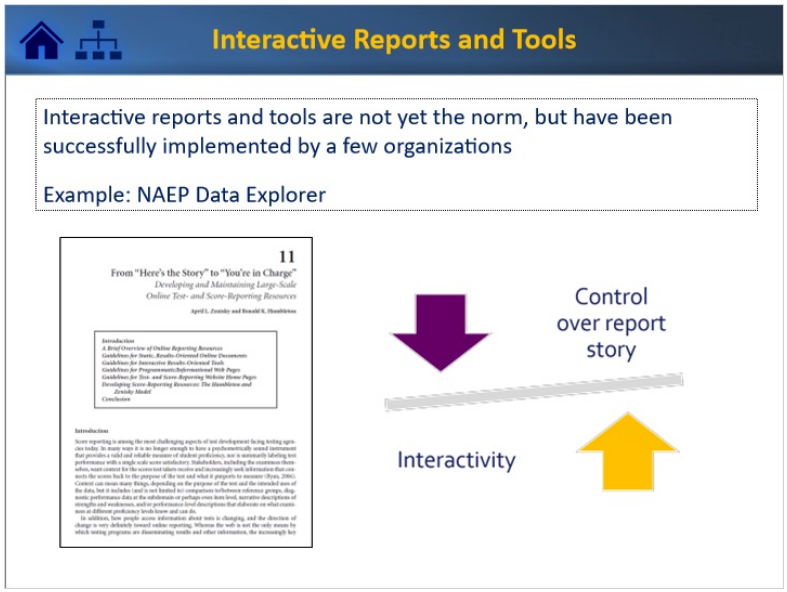
3.27 How I



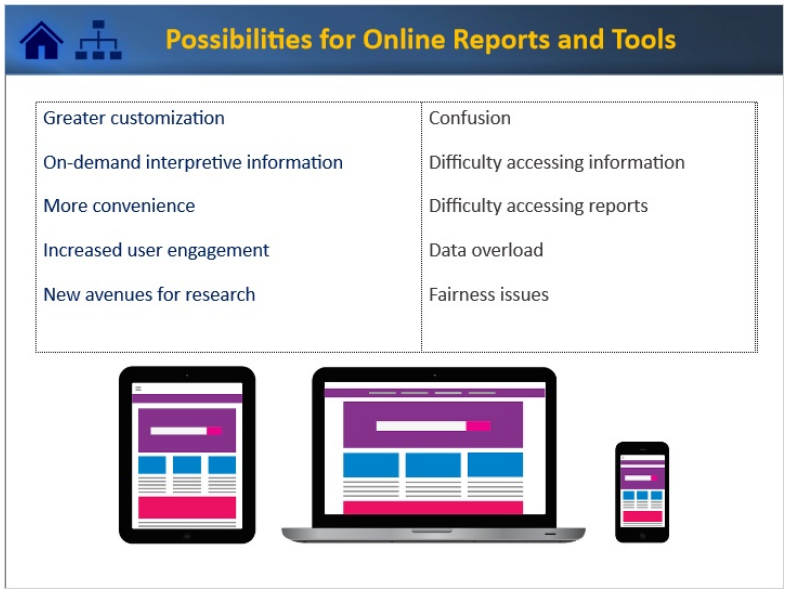
3.28 How II



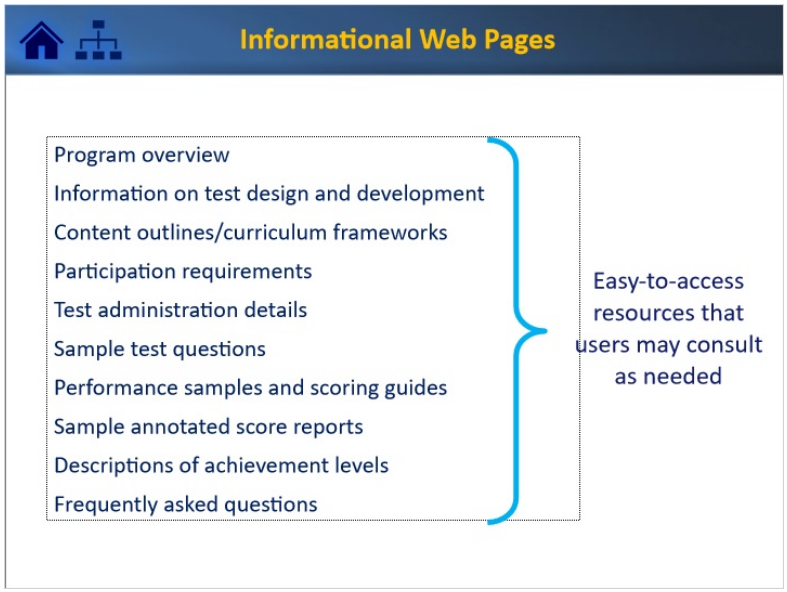
3.29 How III



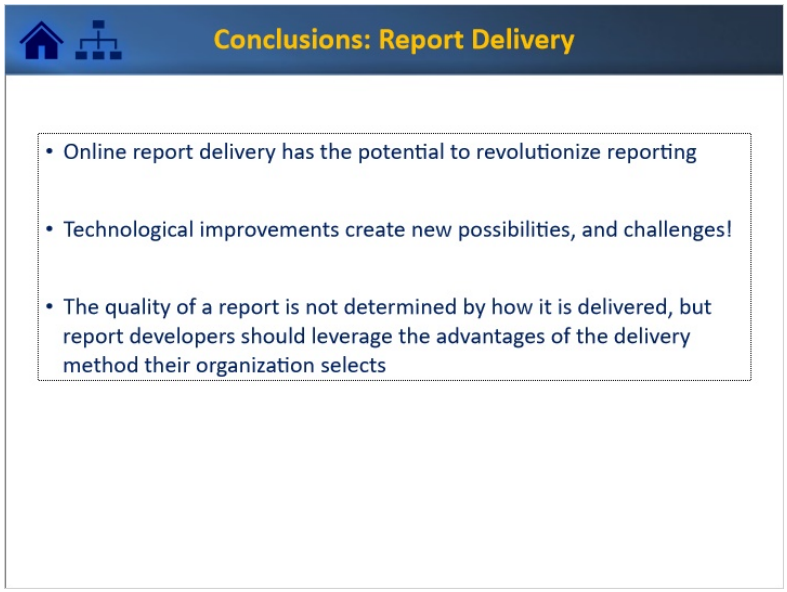
3.30 How IV



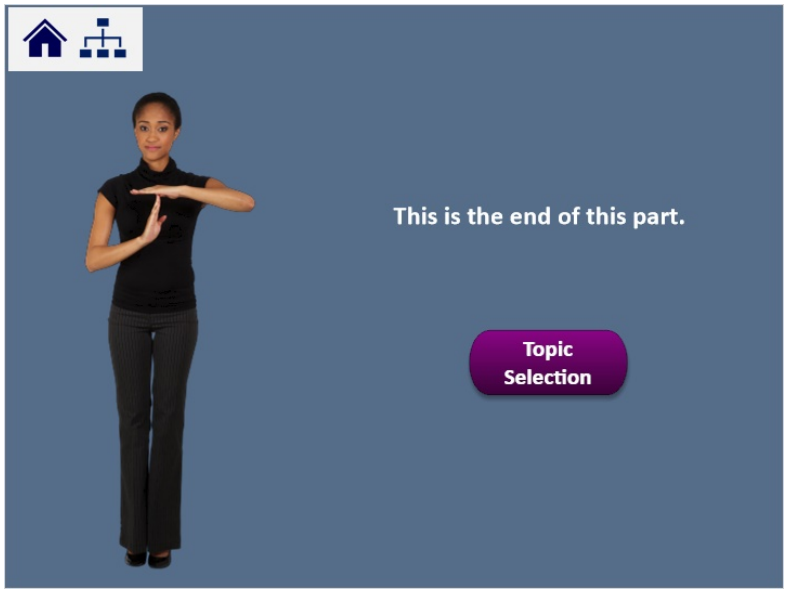
3.31 How V



3.32 How VI



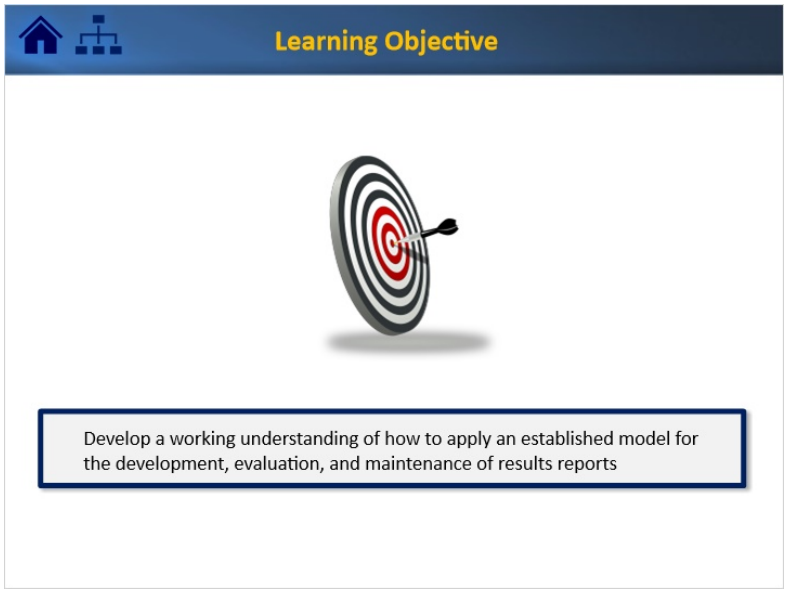
3.33 Bookend: How



4. Section 3: Report Development

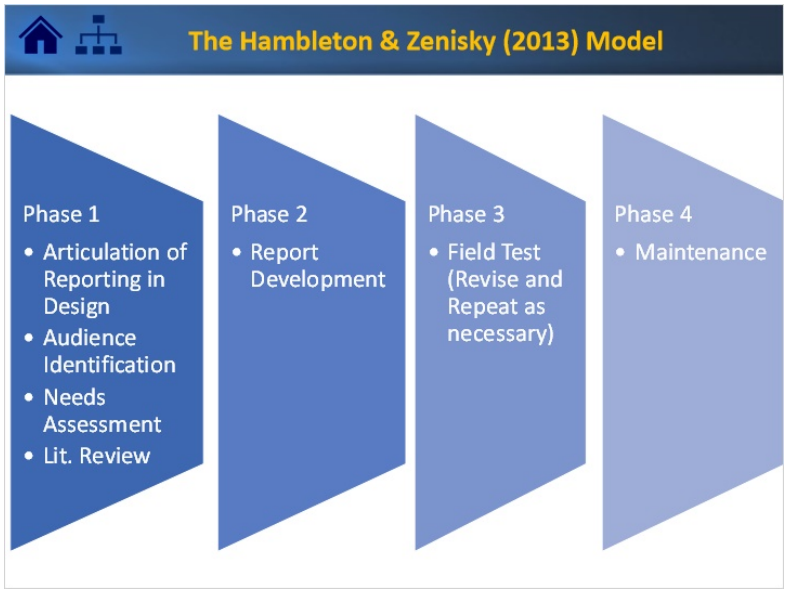
4.1 Cover: Section 3

4.2 Objectives: Section 3

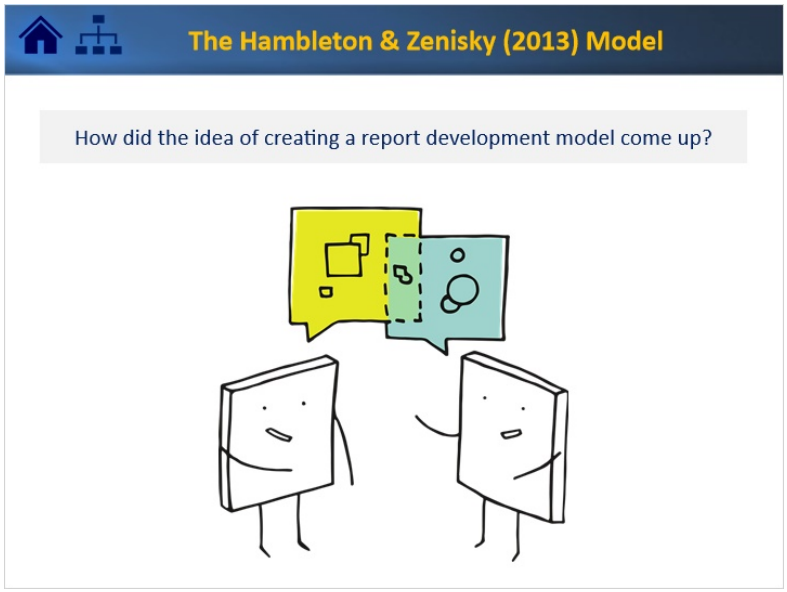


4.3 Interview

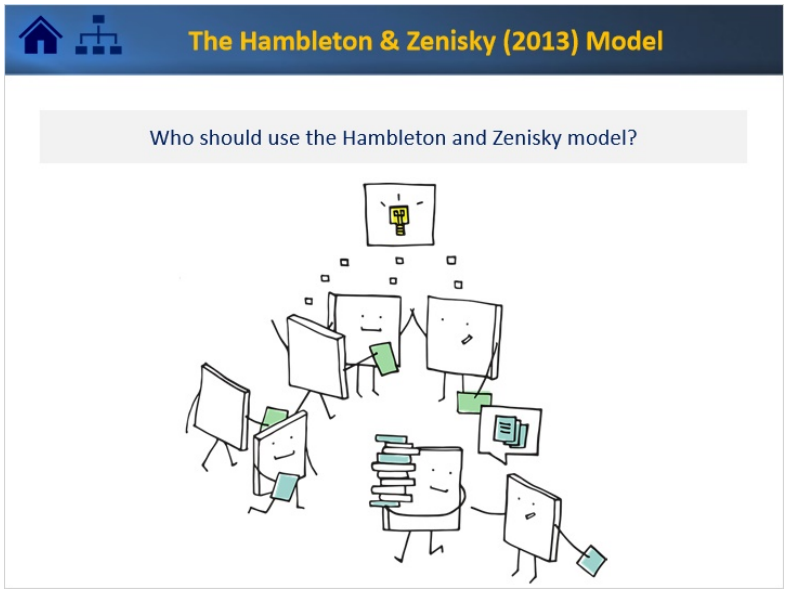
4.4 Model



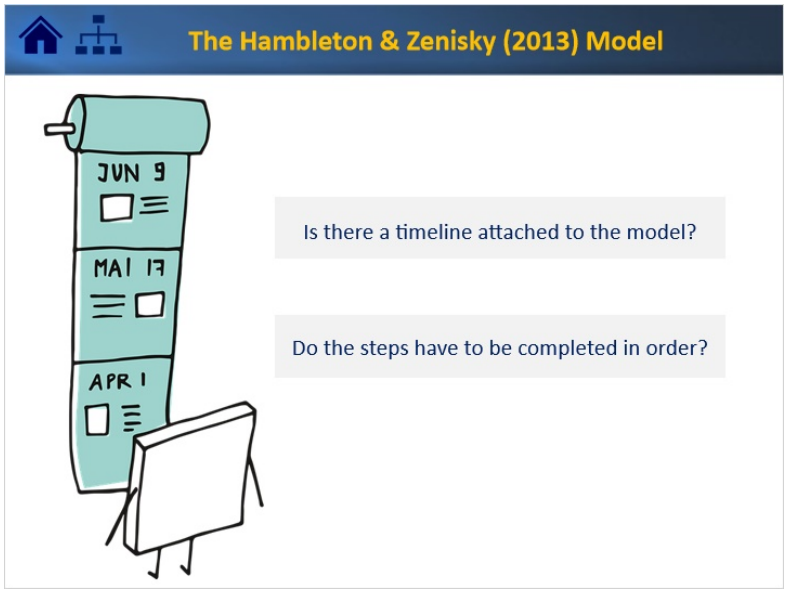
4.5 Model: How



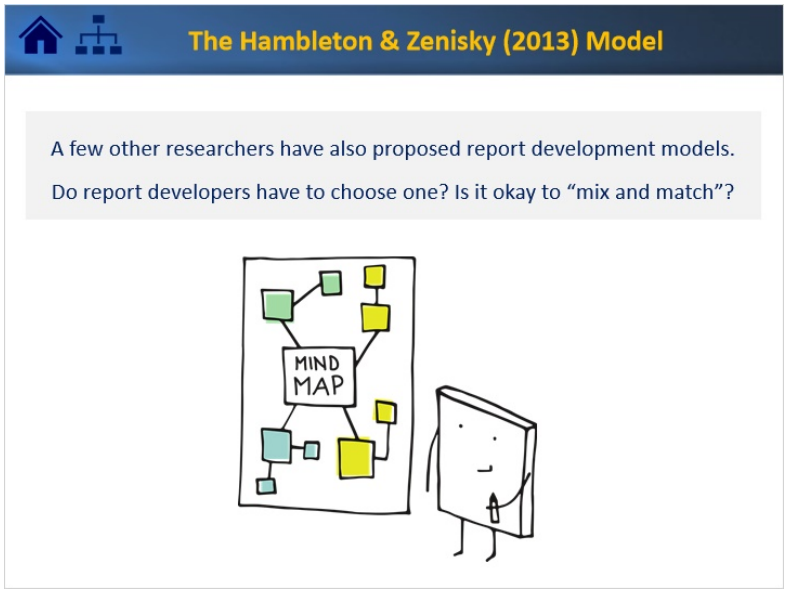
4.6 Model: Who



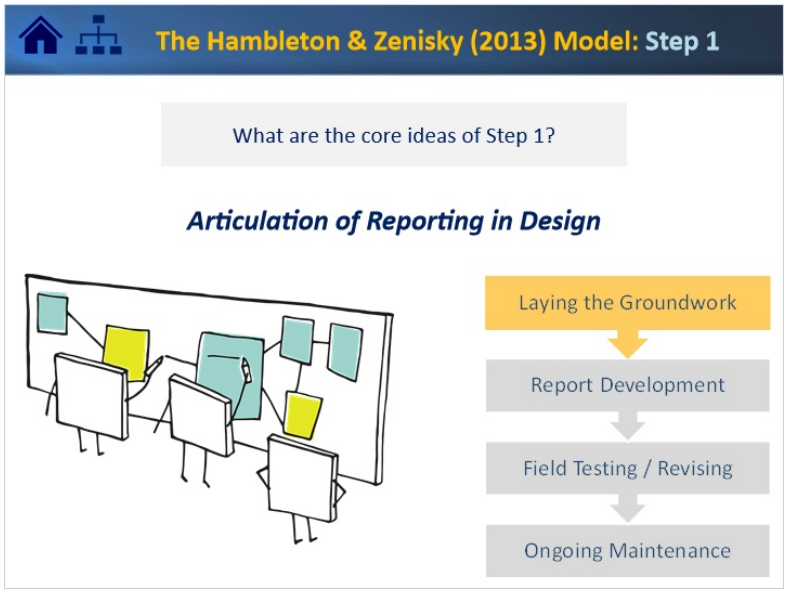
4.7 Model: Timeline



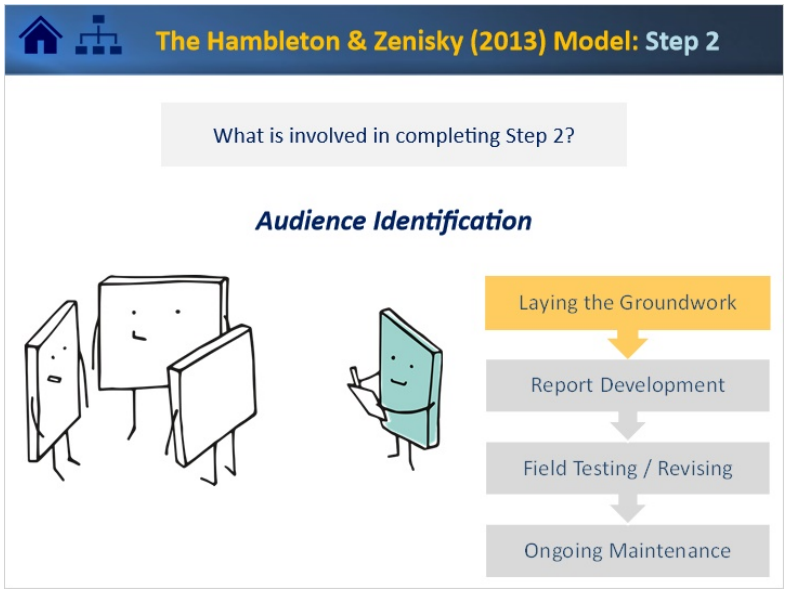
4.8 Model: Mix & Match



4.9 Model: Step 1



4.10 Model: Step 2

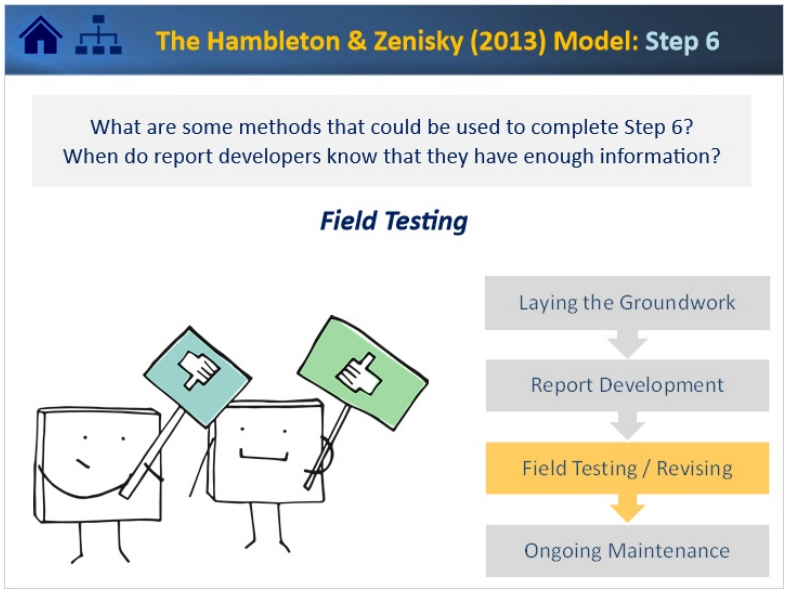


4.11 Model: Step 3

4.12 Model: Step 4

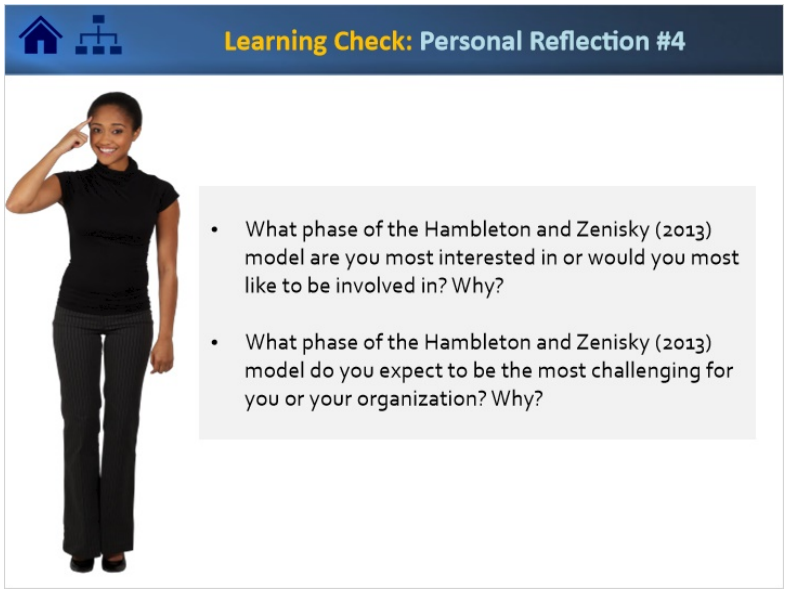
4.13 Model: Step 5

4.14 Model: Step 6

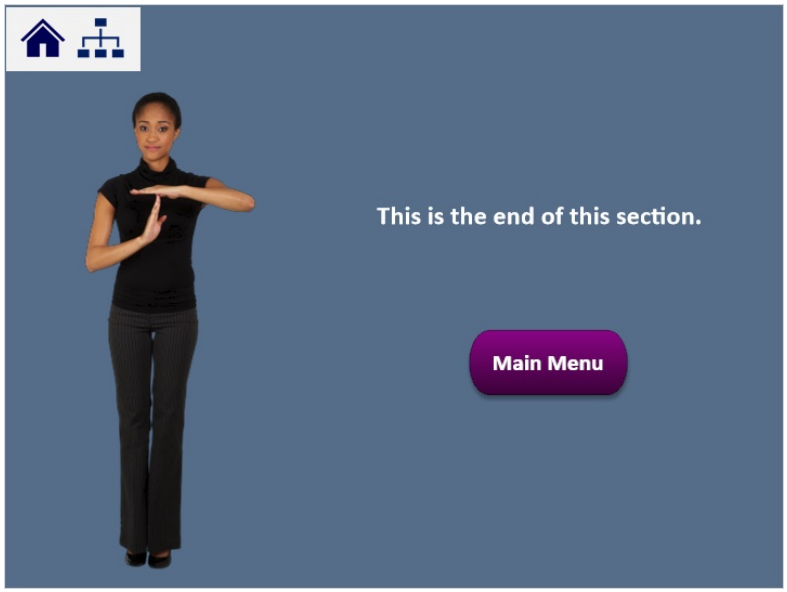


4.15 Model: Step 7

4.16 Learning Check: Personal Reflection #4

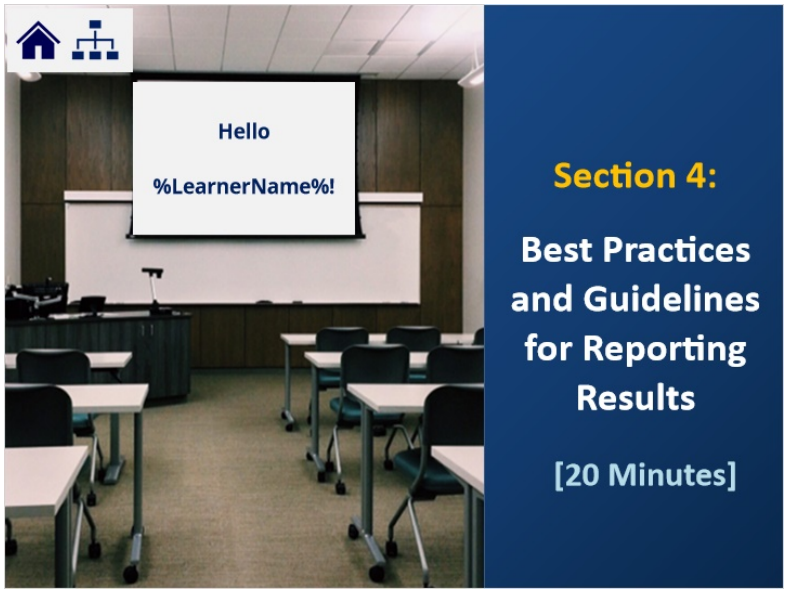


4.17 End of Section 3



5. Section 4: Standards and Guidelines

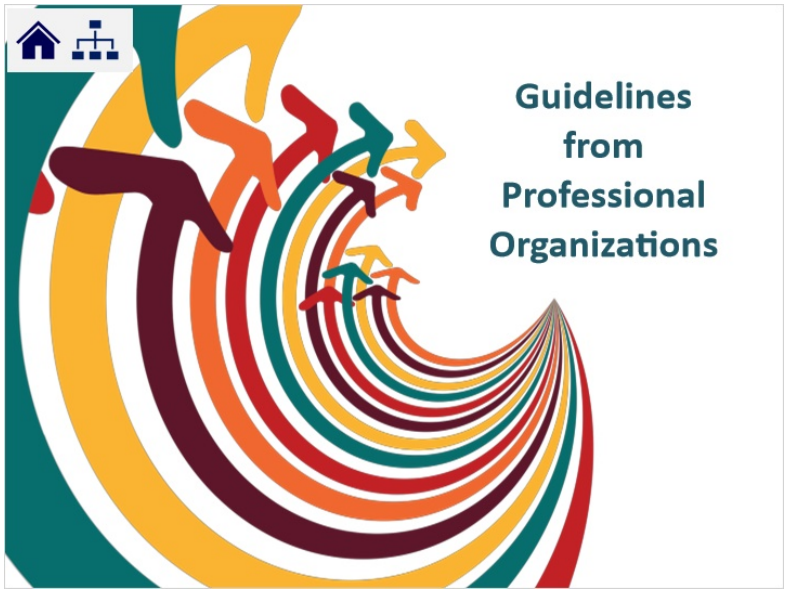
5.1 Cover: Section 4



5.2 Objectives: Section 4

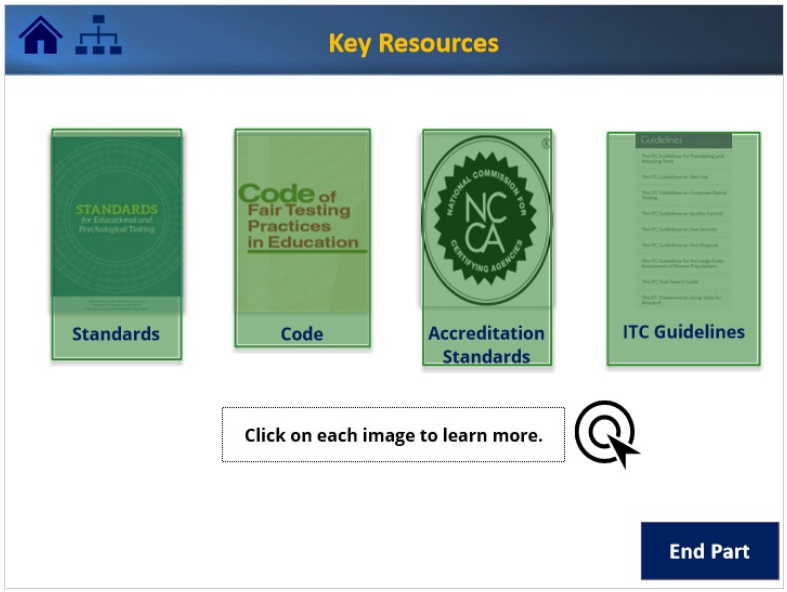
5.3 Topic Selection

5.4 Resource Overview (I)

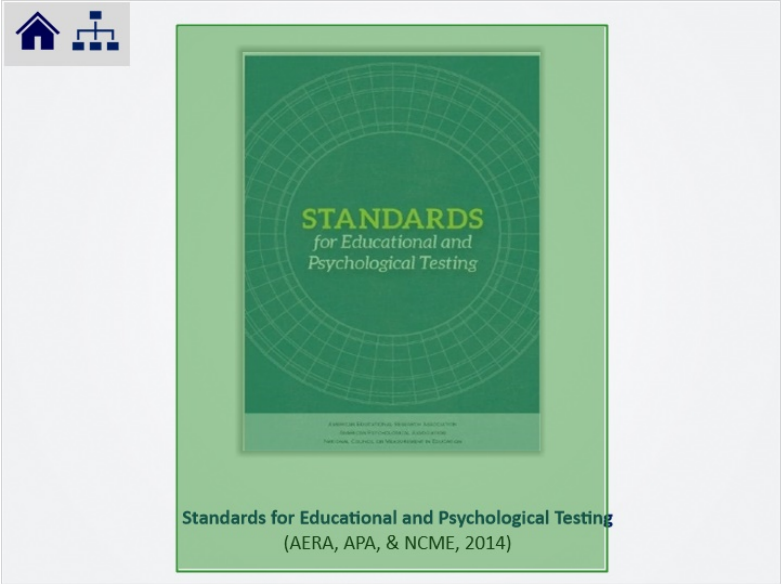


5.5 Resource Overview (I)

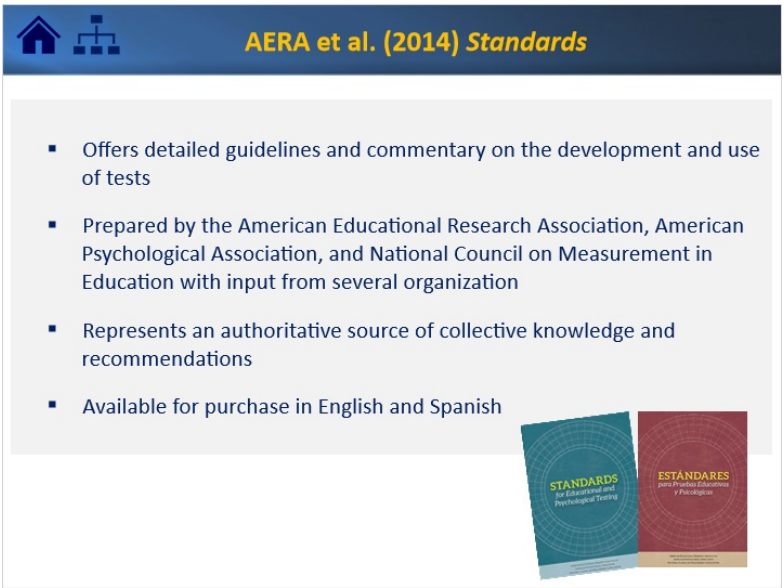
5.6 Resource Selection



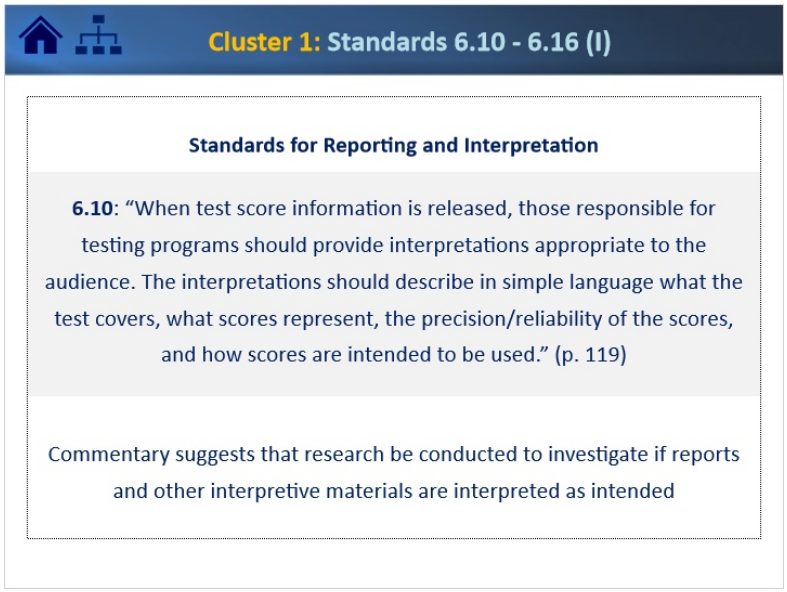
5.7 Bookmark: Joint Standards



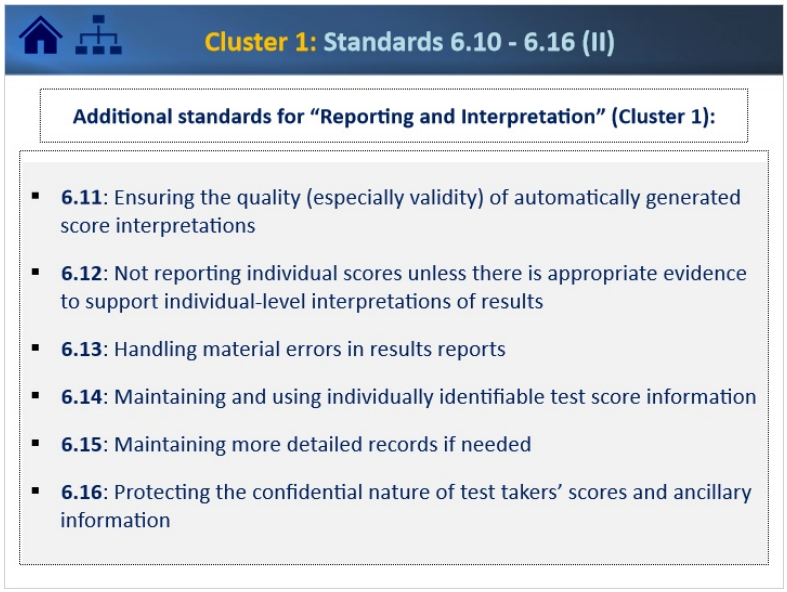
5.8 AERA et al. (2014) Standards



5.9 Standards Clusters



5.10 Additional standards



5.11 Cluster 2: Standards 8.7-8.8

Cluster 2: Standards 8.7-8.8

Standards for "Test Takers' Rights to Fair and Accurate Score Reports"

- **8.7**: Selecting appropriate labels and describing them precisely if scores are assigned to categories (i.e., achievement levels)
  - Appropriate labels are those that "reflect intended inferences" without being "unnecessarily stigmatizing" (p. 136)
- **8.8**: Providing test takers with timely access to reports, unless test takers have waived their right to this information explicitly or implicitly
  - Explicit versus implicit waiver
  - What is considered "timely access"?

5.12 Standards Clusters

Additional Standards

Two clusters of standards…. That's all?

- **No.** All aspects of test development directly or indirectly influence, what information can be reported
- **Several standards in other clusters** provide useful recommendations for report development

5.13 Standard 1.1

Report Development: Standard 1.1

Consider Step 1 of the H&Z report development model and Standard 1.1:

- **Step 1:** Define purpose of report, connect to test purpose and interpretation/use argument.
- **Standard 1.1:** "The test developer should set forth clearly how test scores are intended to be interpreted and consequently used." (p. 23)

5.14 AERA standards

Addressing Misinterpretations: Standard 5.3

Emphasis on safeguarding against possible misinterpretations and misuses of test results:

5.3: "If there is sound reason to believe that specific misinterpretations of a score scale are likely, test users should be explicitly cautioned." (p. 102)

5.15 Recommendations to test users

Information Management

- The *Standards* makes several recommendations to provide certain types of information to test users
- They do NOT require that everything be included in the results report
- They emphasize that any information reported should be appropriate to its intended audience

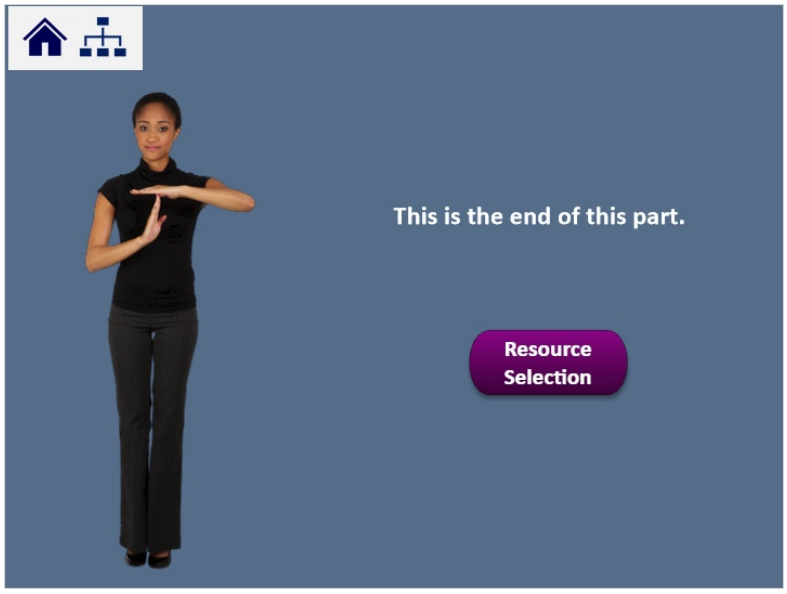
5.16 AERA standard list

List of Applicable AERA et al. (2014) Standards

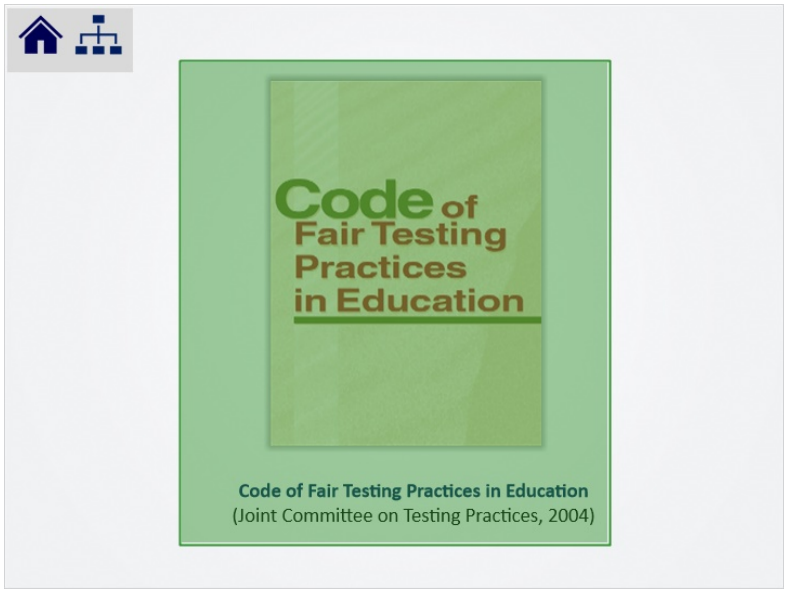
A review of all AERA et al. (2014) standards with recommendations that apply to results reporting is beyond the scope of this module but here is a list of relevant standards to consider for this work:

1.3	2.14	5.10	7.10	11.15
1.5	3.1	6.10	7.11	12.5
1.15	3.17	6.11	8.7	12.11
2.0	4.22	6.12	8.8	12.17
2.3	5.1	6.13	8.10	12.18
2.4	5.3	6.14	9.8	13.5
2.8	5.4	6.15	9.10	13.6
2.10	5.5	6.16	9.16	13.7
2.13	5.8	7.1	11.1	13.9

5.17 Bookend: Joint Standards



5.18 Bookmark: Code

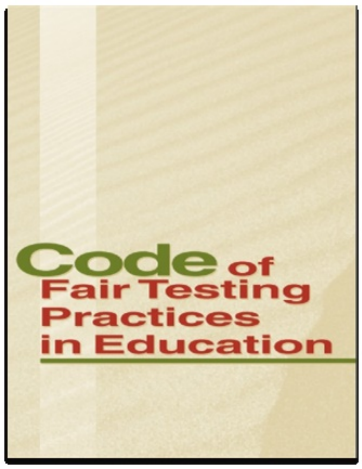


5.19 Code of fair testing

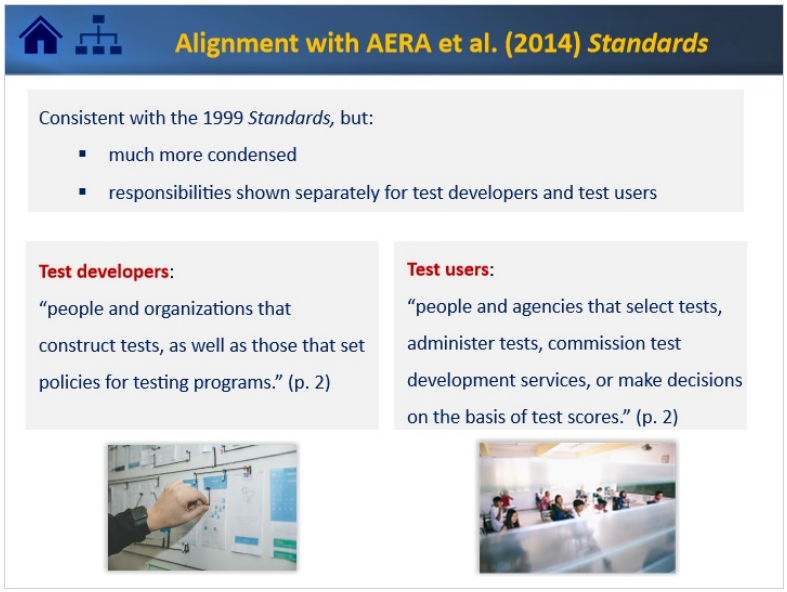
JCTP (2004) *Code*

- Prepared by the Joint Committee on Testing Practices
- Included the same organizations that worked on the AERA et al. *Standards* (2014) and others, such as the National Association of School Psychologists and the National Association of Test Directors
- A practical guide designed to cover testing in education but not employment, licensure, or certification testing
- Available as a free **PDF** on the website of the APA

pdf (Slide Layer)



5.20 Code Alignment



5.21 Code Organization

5.22 Reporting and Intepreting Results

Statements C1 & C3: Connections to Standard 6.10

Topic C: "Reporting and Interpreting Test Results" (C1-C8)

AERA et al. Standard 6.10

When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what scores represent, the precision/ reliability of the scores, and how scores are intended to be used.

Statement C1

Provide information to support recommended interpretations of the results, including the nature of the content, norms or comparison groups, and other technical evidence. Advise test users of the benefits and limitations of test results and their interpretation. Warn against assigning greater precision than is warranted.

Statement C3

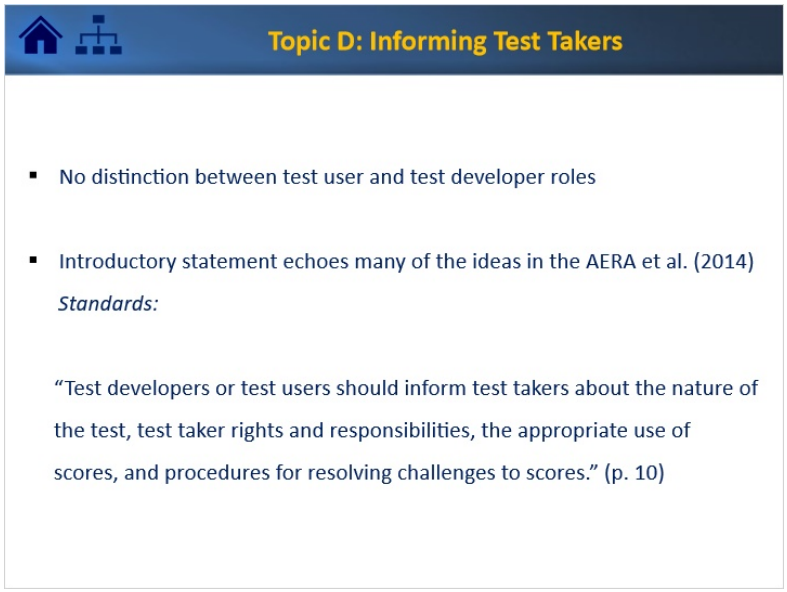
Specify appropriate uses of test results and warn test users of potential misuses.

5.23 Other responsibilities

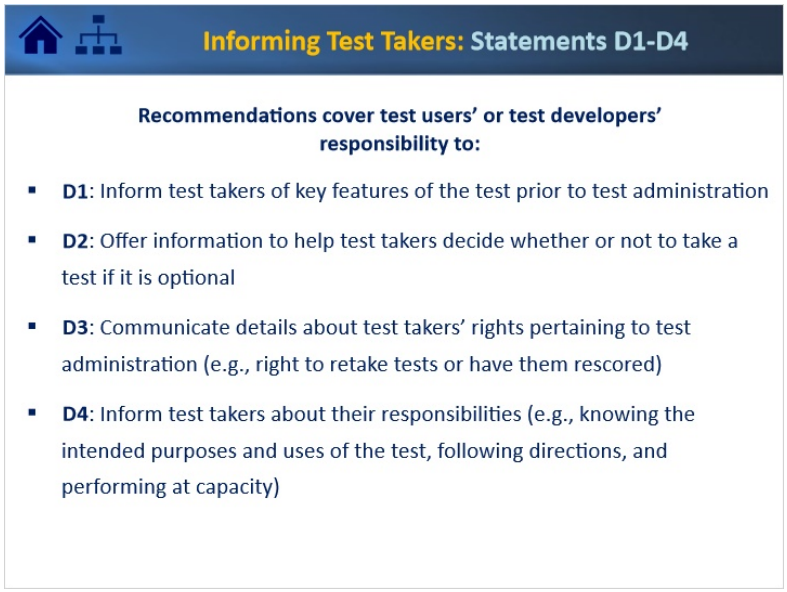
Reporting and Interpreting Test Results

- **C2**: Provide guidance on interpreting results when test are administered with modifications
- **C4**: Present supporting evidence for any performance standards implemented and avoid the use of stigmatizing labels
- **C5**: Encourage test users to consider multiple sources of information to make decisions about a test taker
- **C6**: Provide information to support the interpretation of group-level results
- **C7**: Deliver results in a timely fashion and in a way that is understood by test takers
- **C8**: Offer guidance to test users on how to monitor the degree to which the test is fulfilling its intended purpose

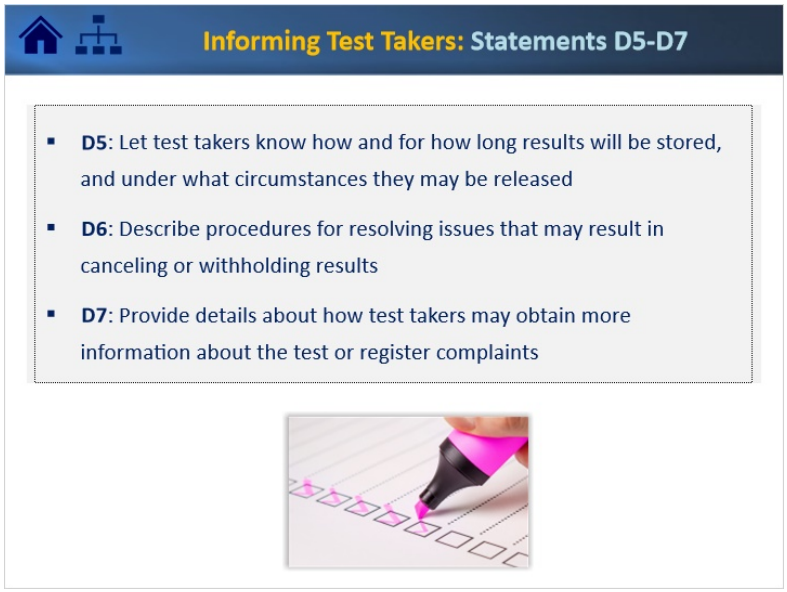
5.24 Informing Test Takers



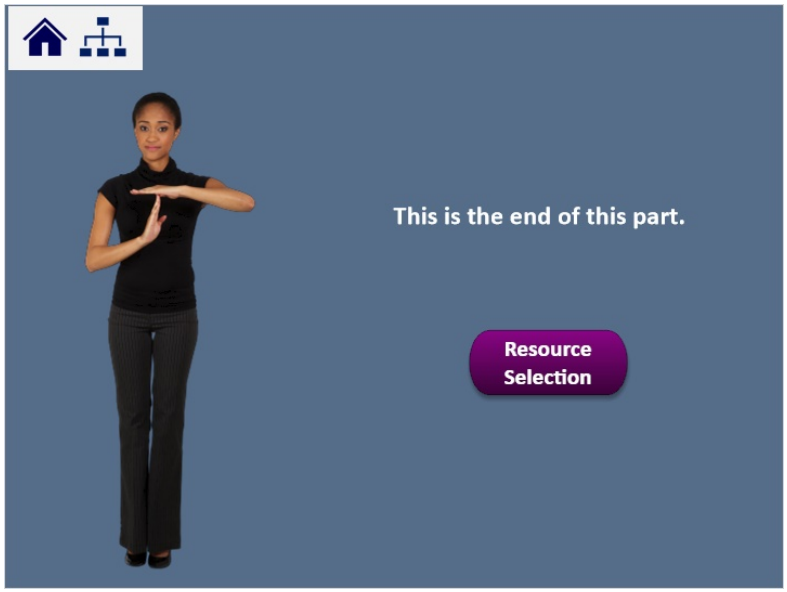
5.25 Recommendations for test user and test developer



5.26 Recommendation Continued

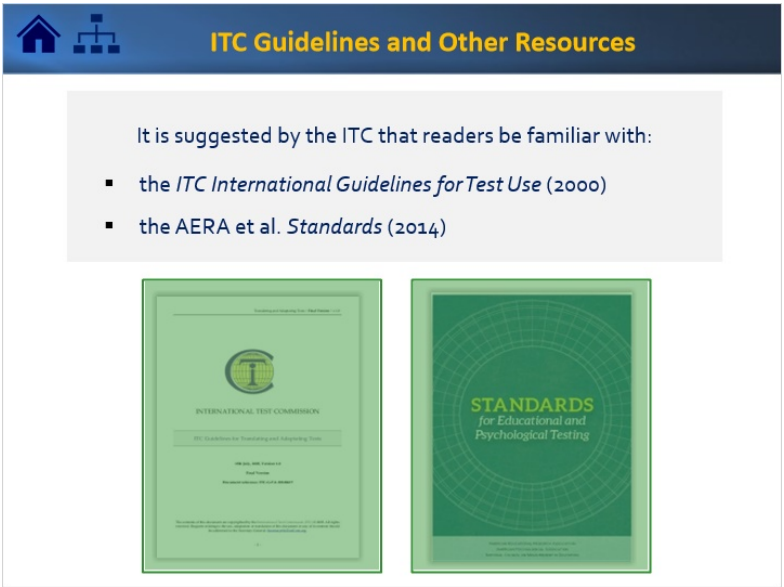


5.27 Bookend: Code

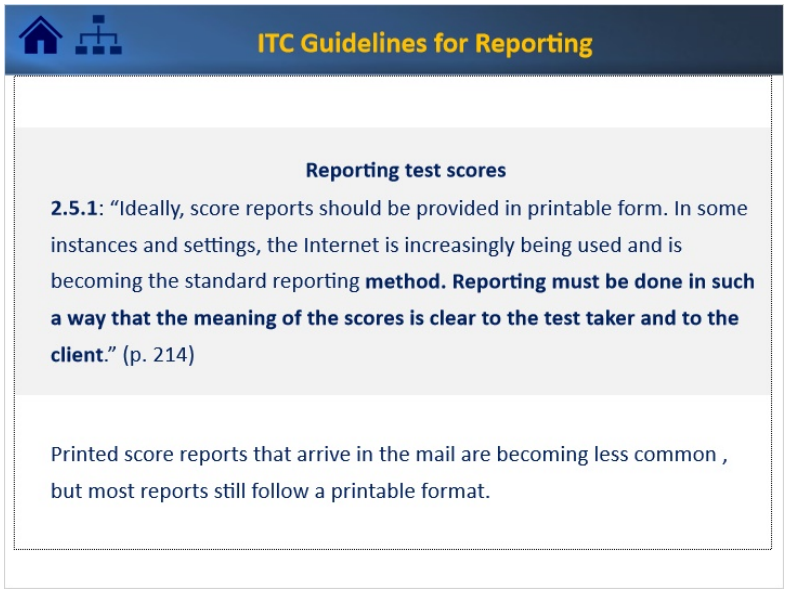


5.28 Bookmark: ITC Guidelines

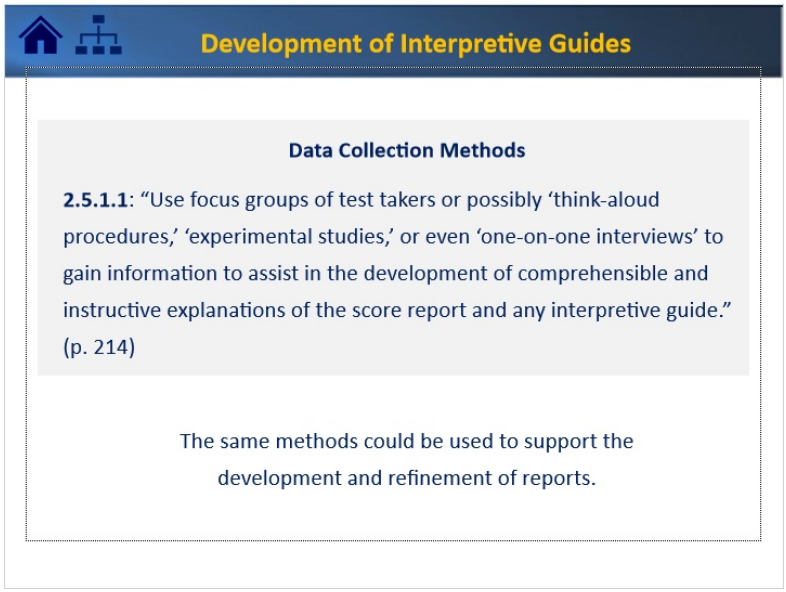
5.29 ITC Alignment



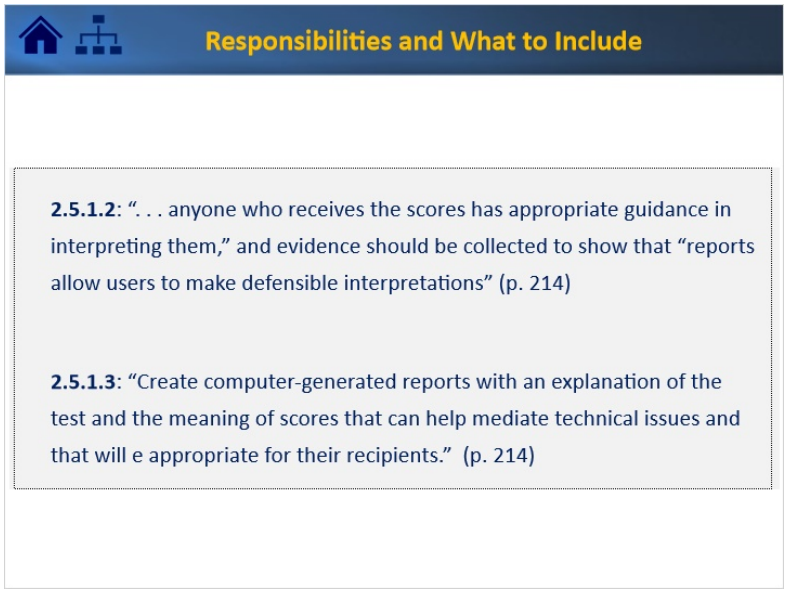
5.30 Clusters pertaining results reporting



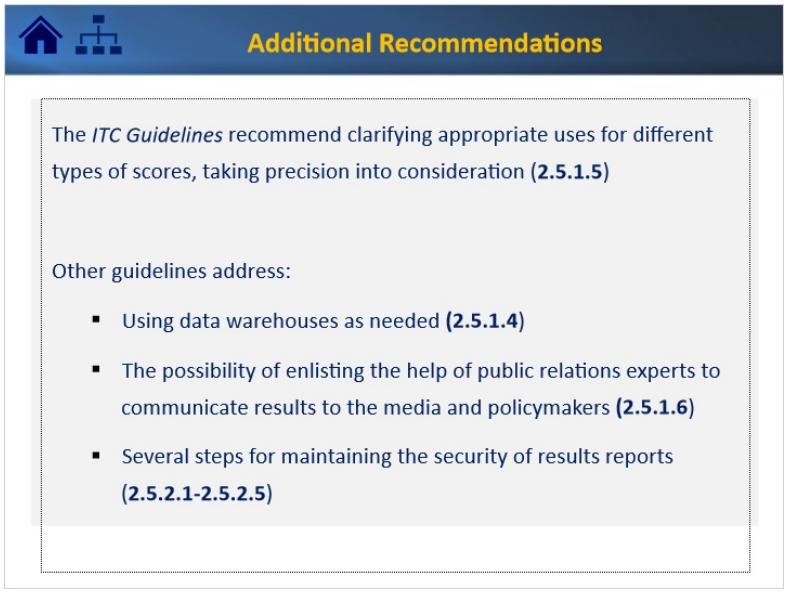
5.31 Example of methods



5.32 Guidelines of responsibilities



5.33 ITC Guidelines recommendations



5.34 Bringing it all together

Bringing it All Together: Conclusions

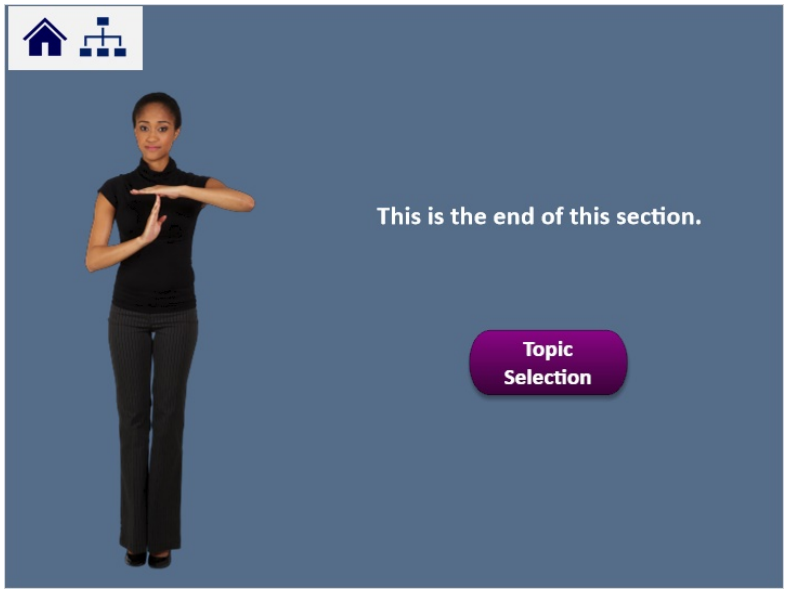
- The AERA et al. (2014) *Standards* should be **"required reading"**
- Other sources (*JCTP Code*, *NCCA Standards, and ITC Guidelines*) also provide valuable information and may be required reading **depending on one's field**
- None of the sources provide step-by-step instructions for developing a perfect report (there is **no one-size-fits-all solution!**)
- All sources provide key recommendations for developing reports that **support appropriate interpretations and uses** and diminish the likelihood of misinterpretations and misuses
- We can turn to the **research literature** for additional recommendations

5.35 Learning check: Reflection

Learning Check: Personal Reflection #5

In reflecting on the various standards presented, how do you see yourself using each of the resources - or not using them - in your reporting efforts in the future?

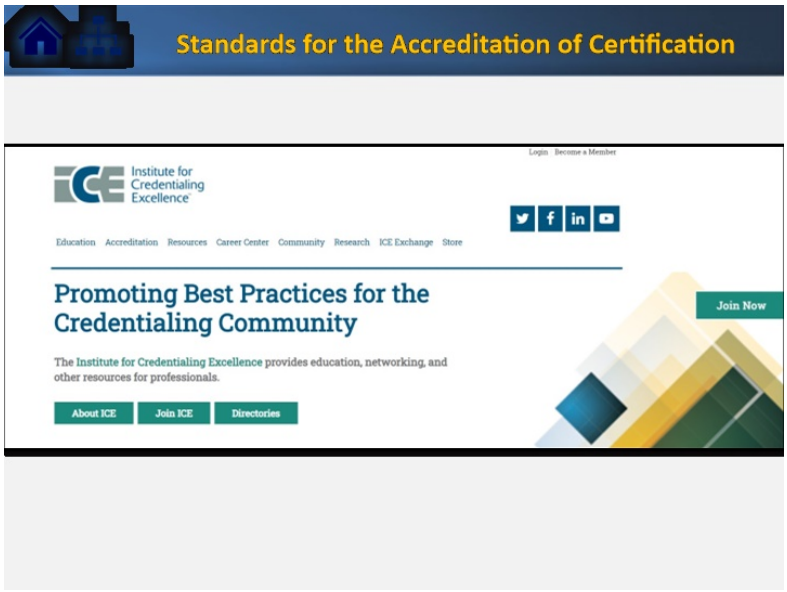
5.36 Bookend: ITC Guidelines



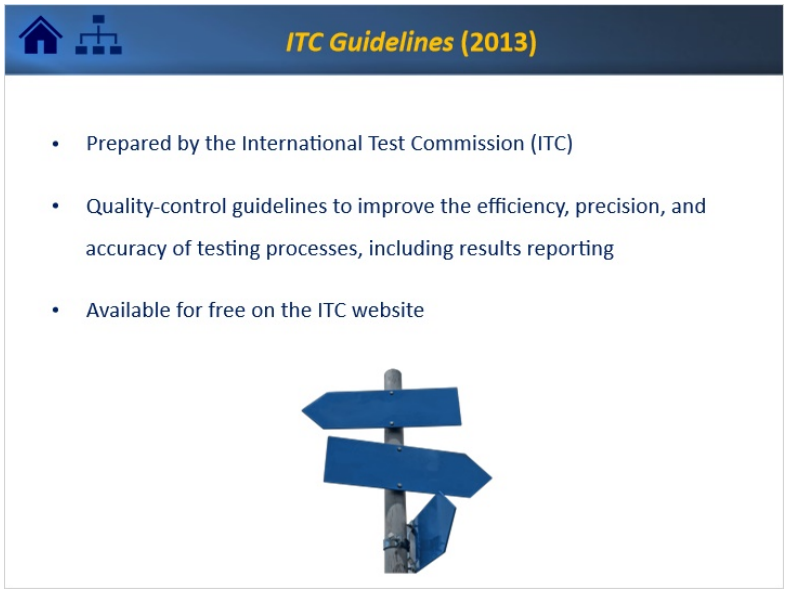
5.37 Bookmark: Accreditation Standards

5.38 Standards for the Accreditation of Certification Programs (2014)

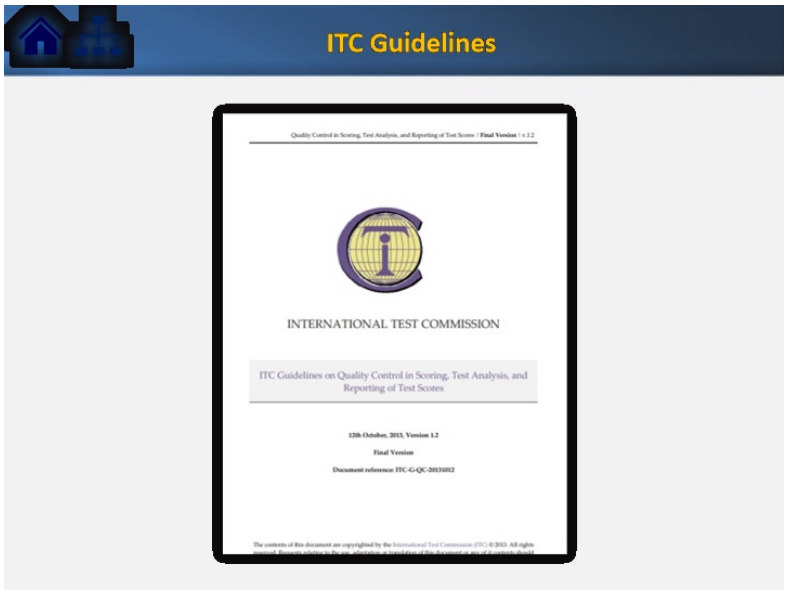
website (Slide Layer)



5.39 ITC Guidelines



Pdf (Slide Layer)



5.40 Alignment NCCA

Connections

- The *NCCA Standards* (2014) are consistent with the AERA et al. *Standards* (2014), but focus on the certification context
- Like the AERA et al. *Standards,* the *NCCA Standards* offers guidelines ("essential elements") along with commentary

5.41 NCCA Standards

Essential Elements D & E (Standard 19)

In certification, differential reporting based on whether examinees passed or did not pass is a sensible possibility

D: "All candidates must be provided information on their overall performance on an examination." (p. 26)

E: "Failing candidates must be provided with information about their performance in relation to the passing standard. If the program provides feedback to candidates such as domain-level information, candidates must be provided guidance about limitations in interpreting and using that feedback." (p. 26)

5.42 NCCA Standards 19

Commentary on Reporting Responsibilities

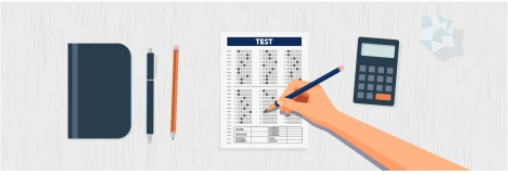
"The certification program should provide candidates with an explanation of the types of scores reported, appropriate uses, and potential misuses" (p. 26)

"If domain-level information has low reliability, programs are advised against reporting it to candidates and other stakeholders. When domain-level or other specific feedback is given to candidates, the certification program should provide estimates of its precision and/or other guidance." (p. 26)

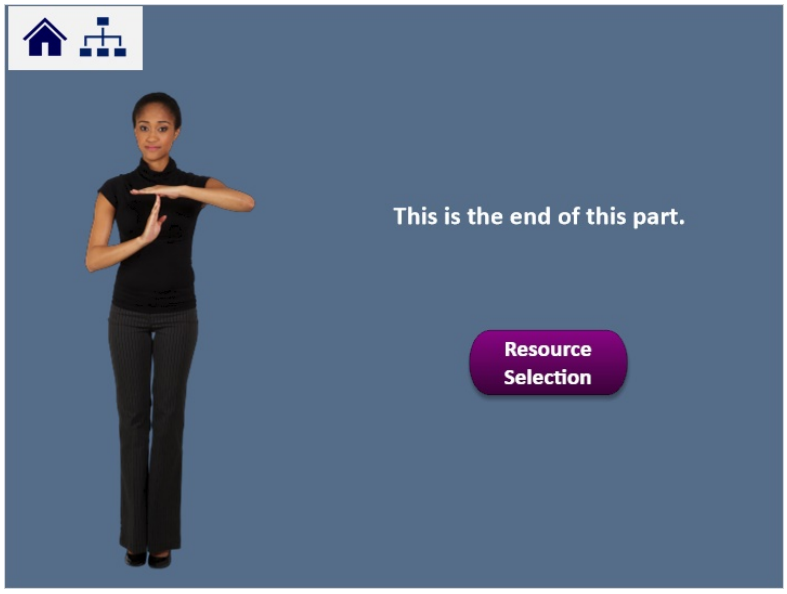
5.43 NCCA standards II

Procedural Aspects of Reporting (Standards 9-10)

- **Standard 9**: How long to maintain examinees' test results and other certification-related records
- **Standard 10**: How to adequately protect the confidentiality of examinees' results and related information

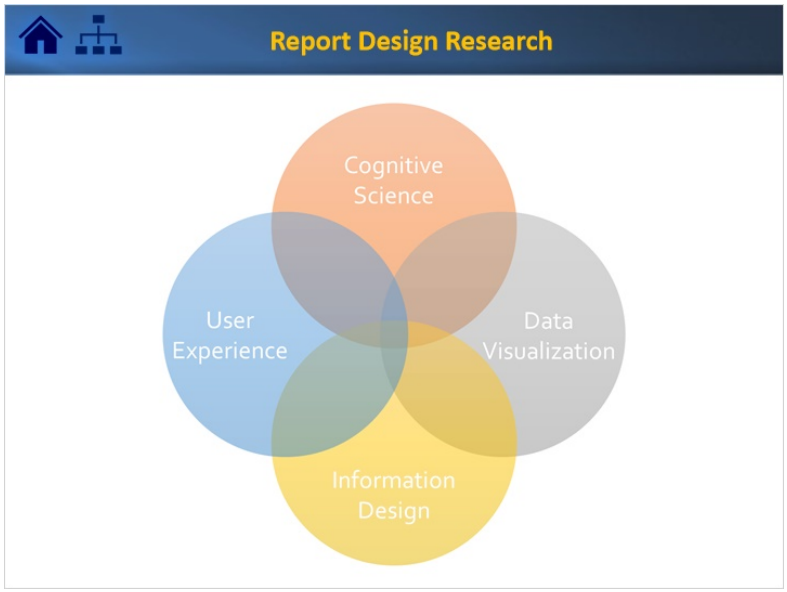


5.44 Bookend: Accreditation Standards



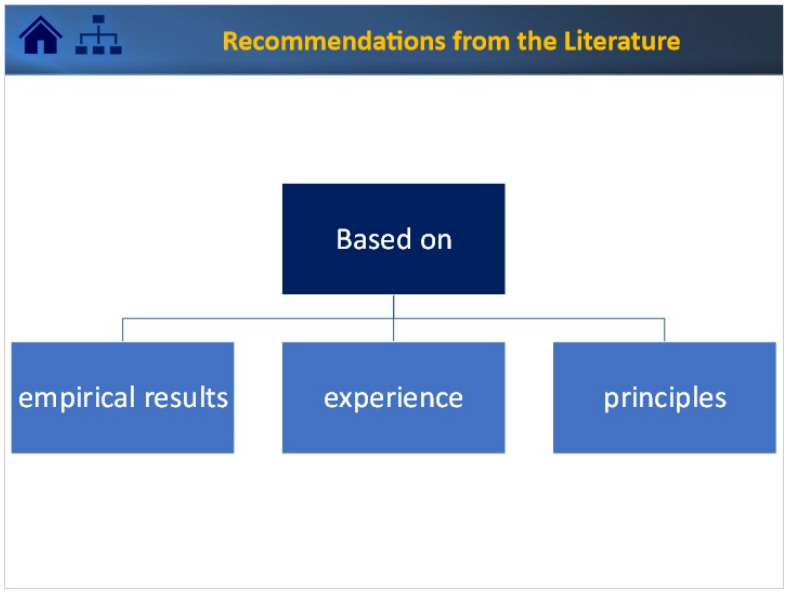
5.45 Bookmark:Best Practices for Reporting Results

5.46 Report Design Research



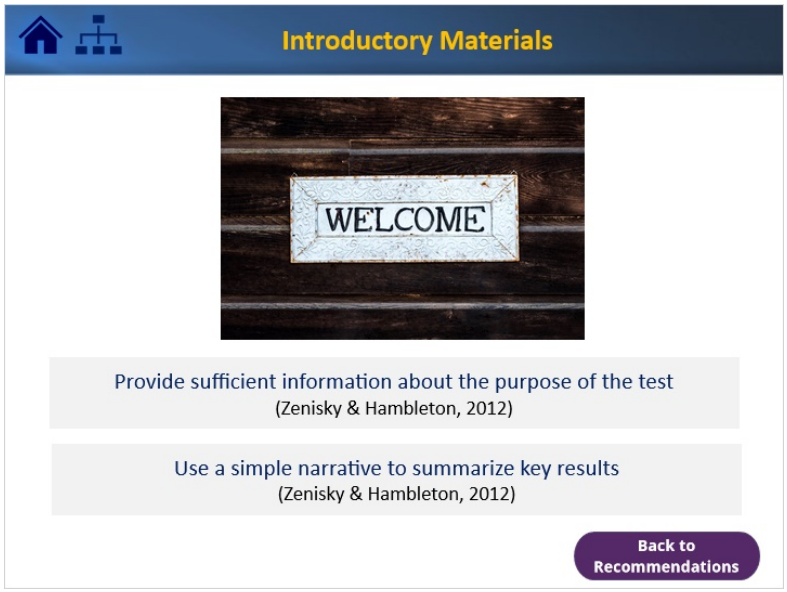
5.47 Generalizability

5.48 Recommendations from Literature

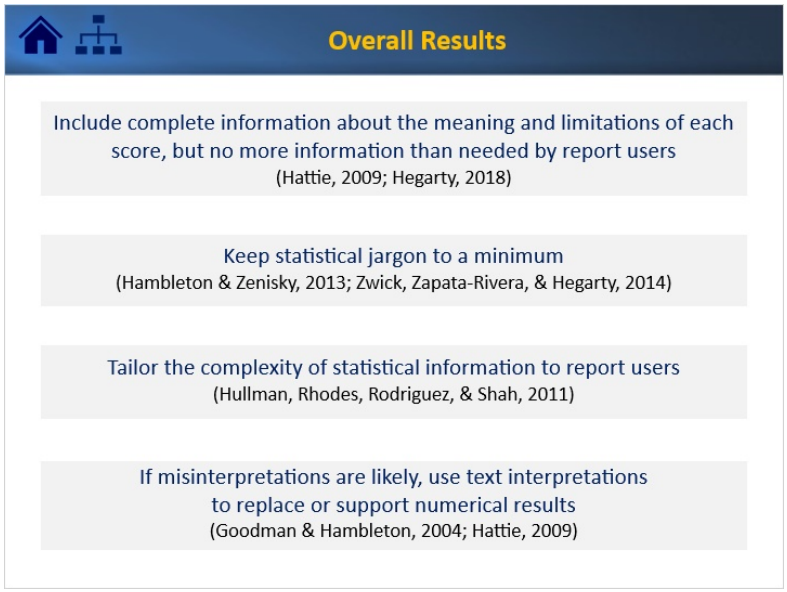


5.49 Topic Selection

5.50 Introductory Materials



5.51 Overall Results I



5.52 Overall Results: Graphics

Overall Results: Graphics

Ensure that text and graphic elements are sufficiently large and consistent with data visualization principles
(Hegarty, 2018; Hullman et al., 2011; Slater, Livingston, & Silver, 2018)

Use graphics that are clear, simple, and satisfy specific purposes
(Goodman & Hambleton, 2004; Ryan, 2006; Wainer, 1997)

Do not rely too heavily on legends and interpretive guides
(Hegarty, 2018)

Use text to support the interpretation of graphics
(Goodman & Hambleton, 2004; Ryan, 2006)

5.53 Overall Results: Caution

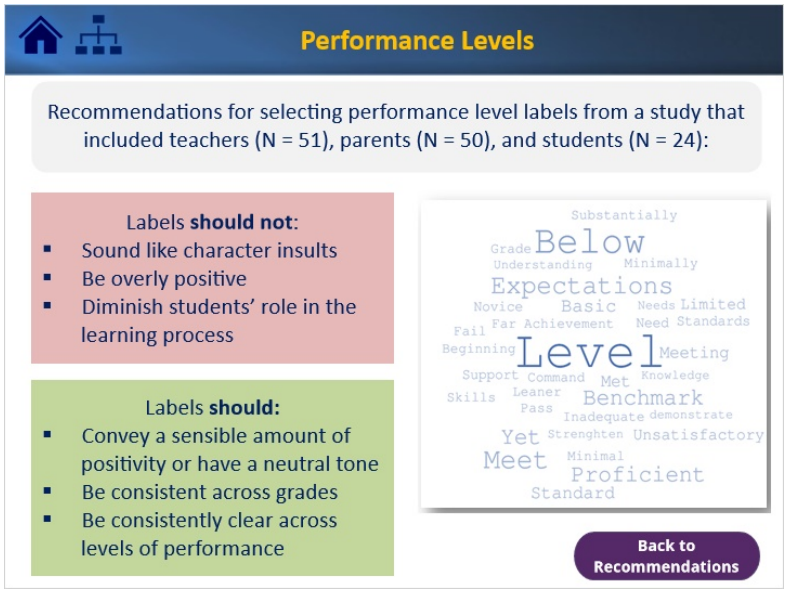
Overall Results: Precision

Be aware that report users prefer different representations of error depending on their level of comfort with statistics
(Zwick et al., 2014)

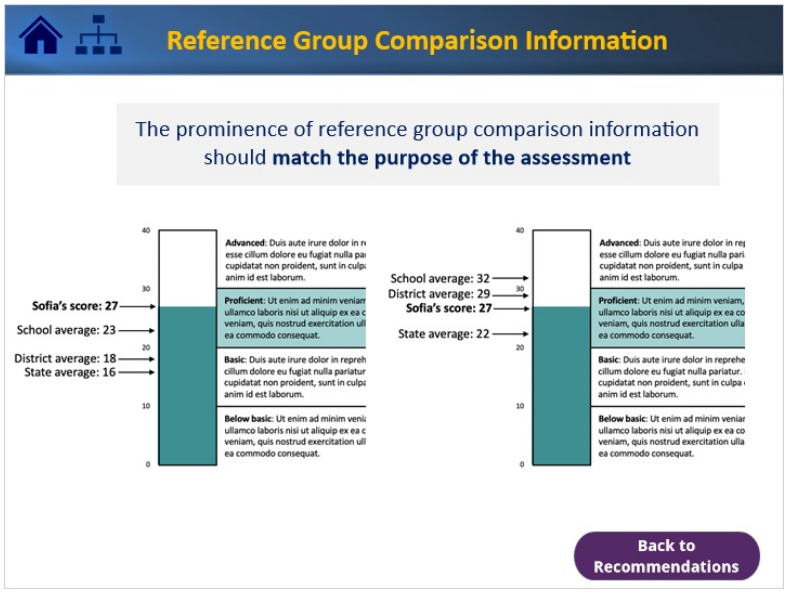
The question of whether information about measurement error is beneficial to report users has no definitive answer
(Kannan, Bryant, Zapata-Rivera, & Peters, 2017;
Kannan, Zapata-Rivera, & Leibowitz, 2018; Zapata-Rivera, Kannan, & Zwick, 2018)

Back to Recommendations

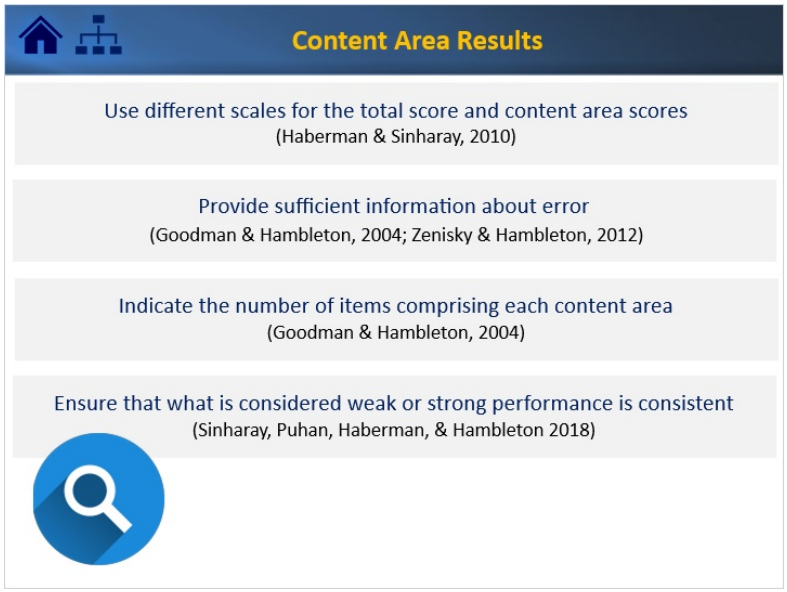
5.54 Performance Level



5.55 Reference group comparison information

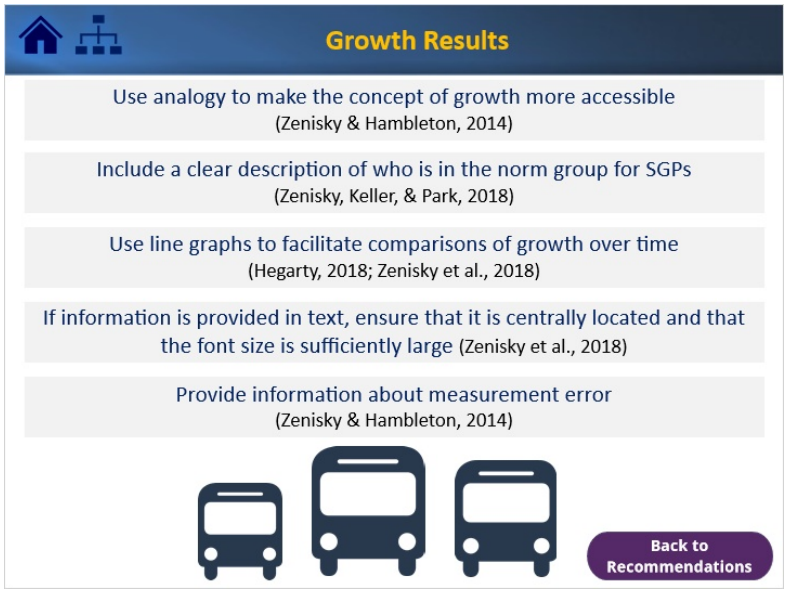


5.56 Content area results

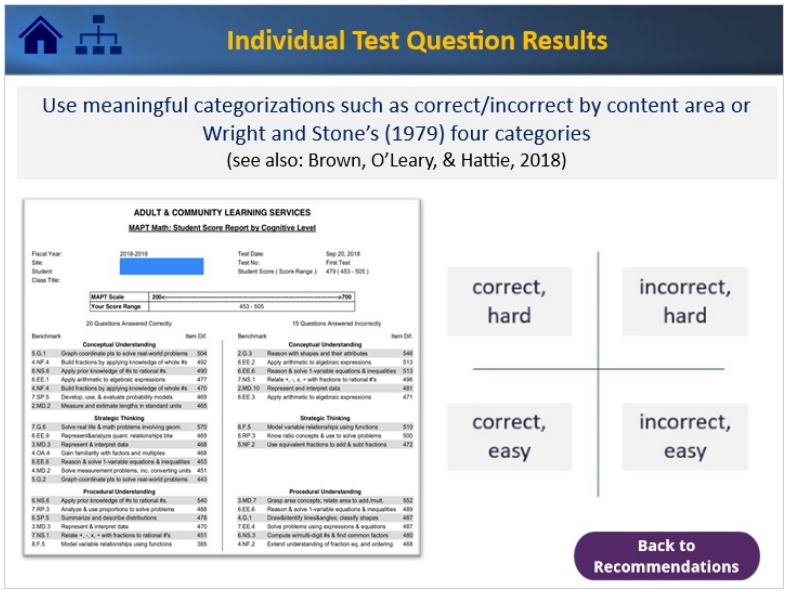


5.57 Content area results II

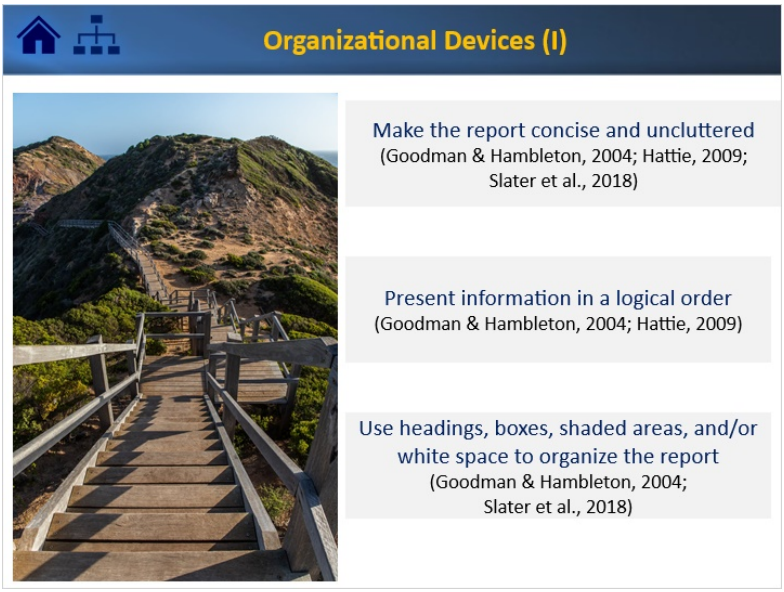
5.58 Growth results



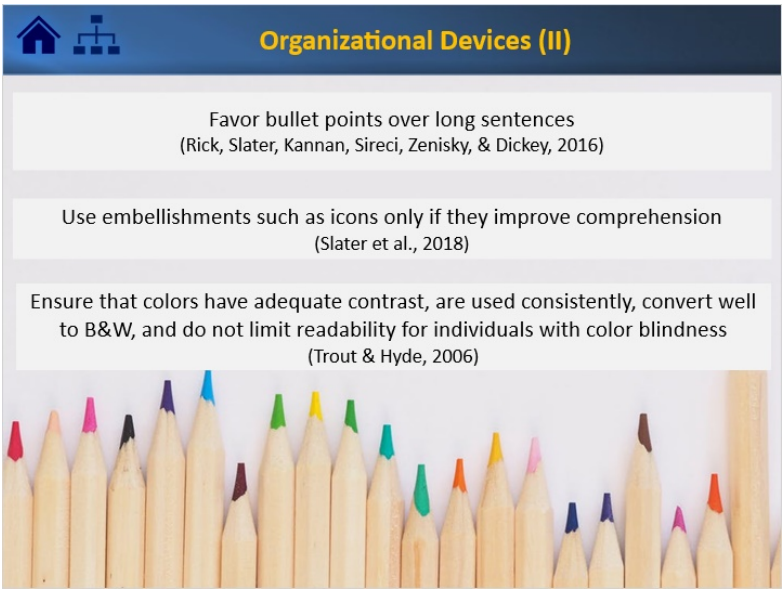
5.59 Individual test question results



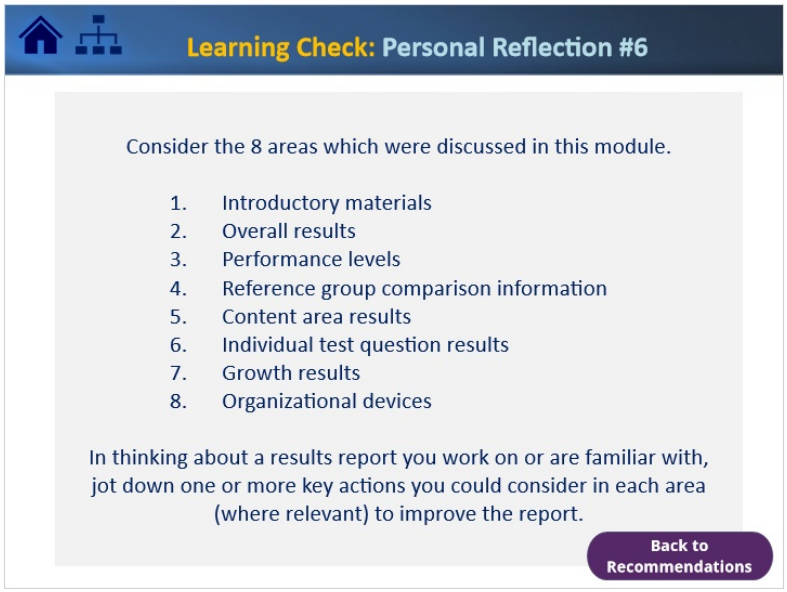
5.60 Organizational devices



5.61 Organizational devices II



5.62 Learning Check: Personal Reflection #6

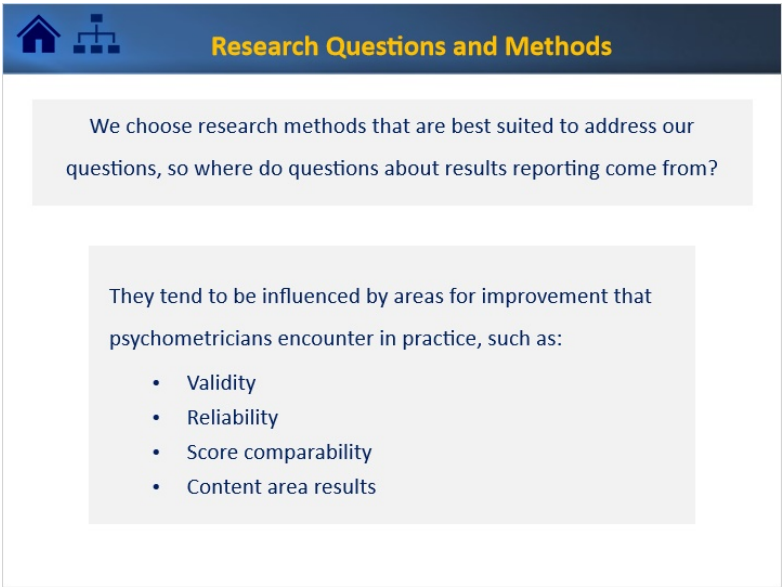


6. Section 5: Research on Reporting

6.1 Cover: Section 5

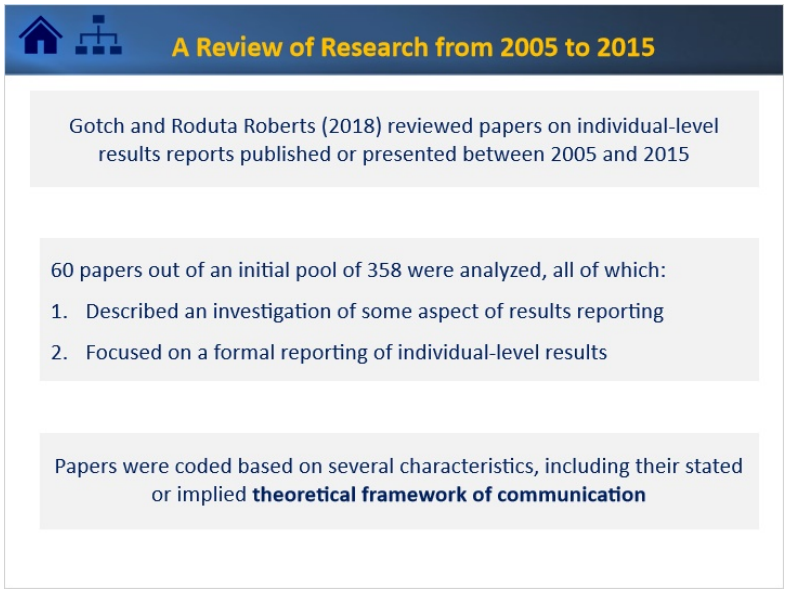
6.2 Objectives: Section 5

6.3 Research Questions and Methods

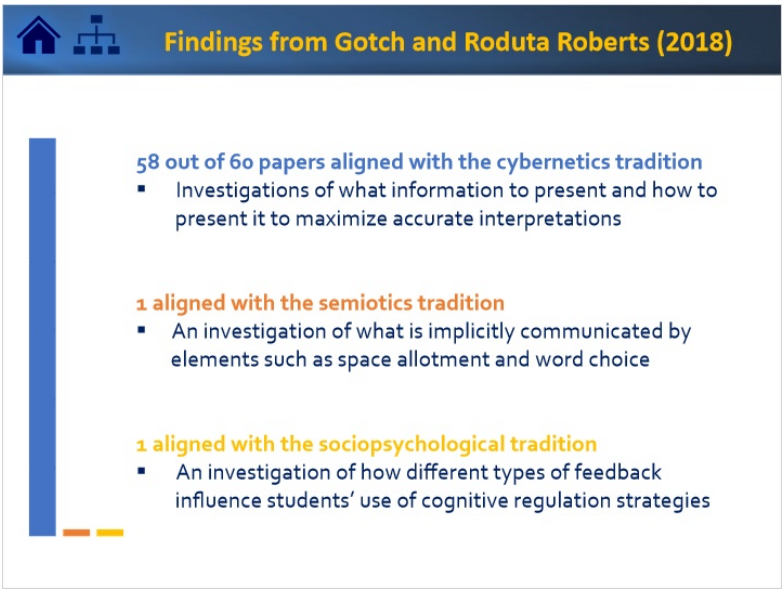


6.4 Areas of focus

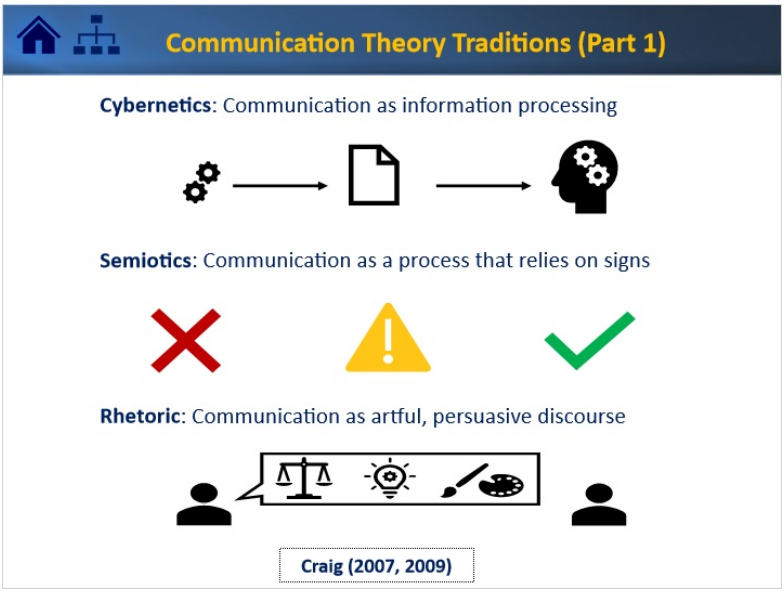
6.5 Areas of focus



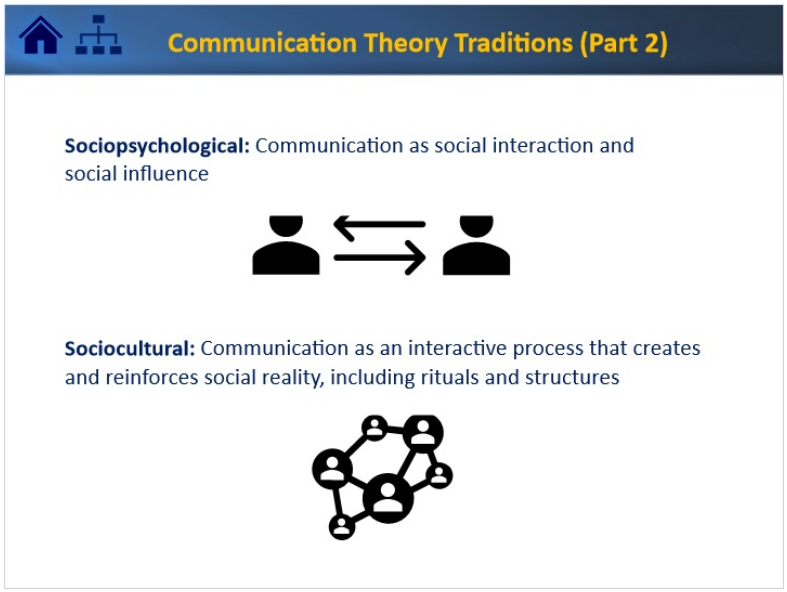
6.6 Findings



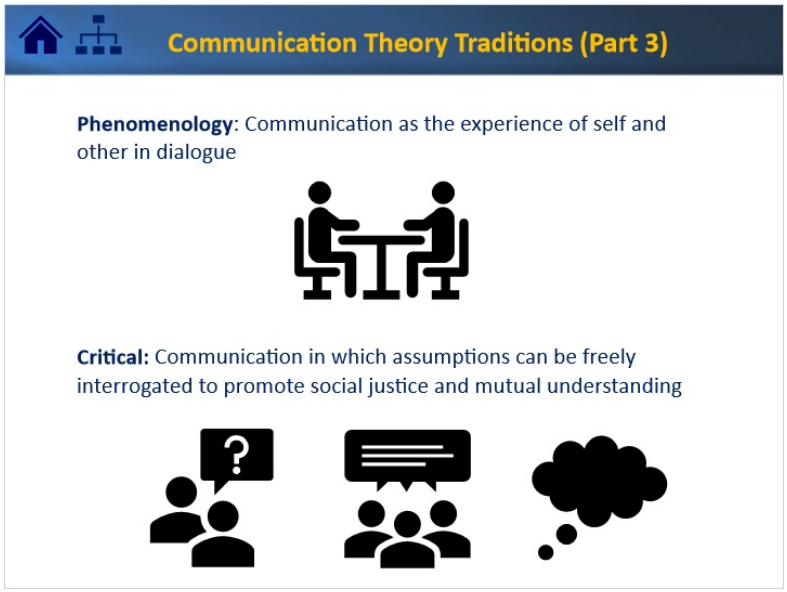
6.7 Communication theory traditions (part 1)



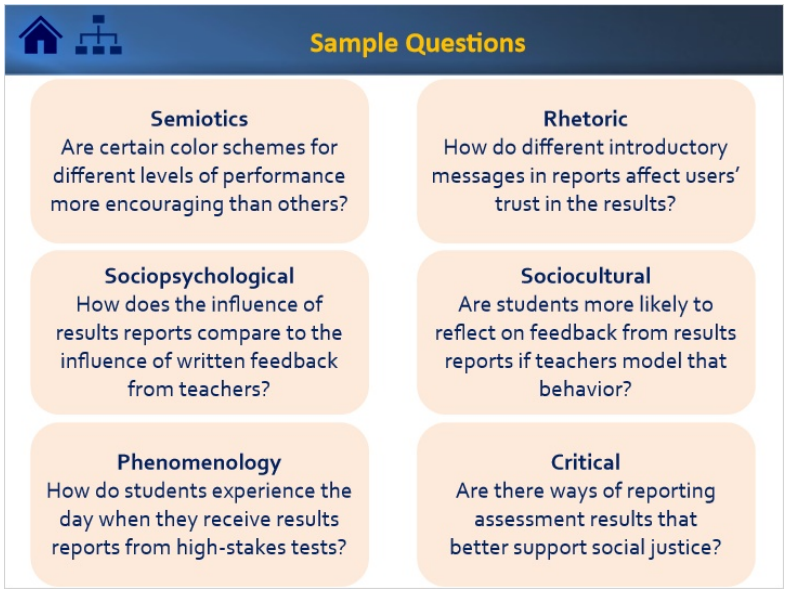
6.8 Communication theory traditions (part 2)



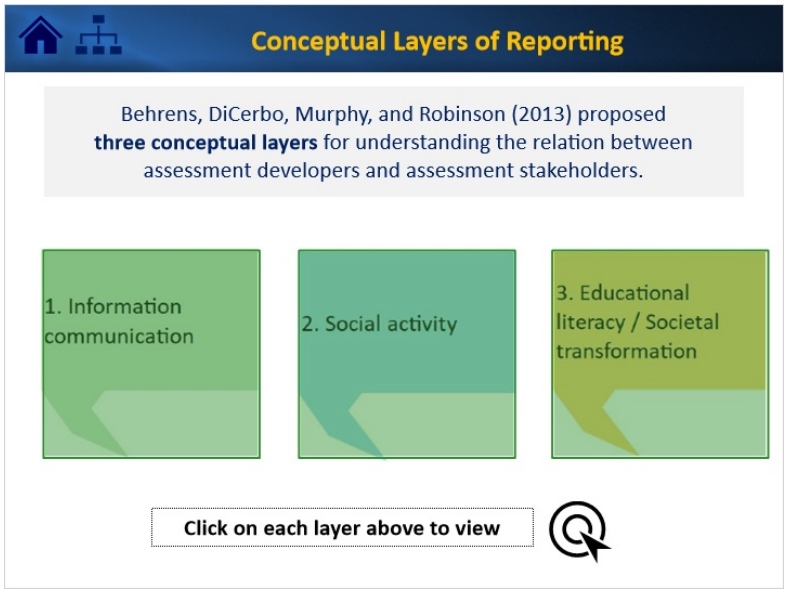
6.9 Communication theory traditions (part 3)



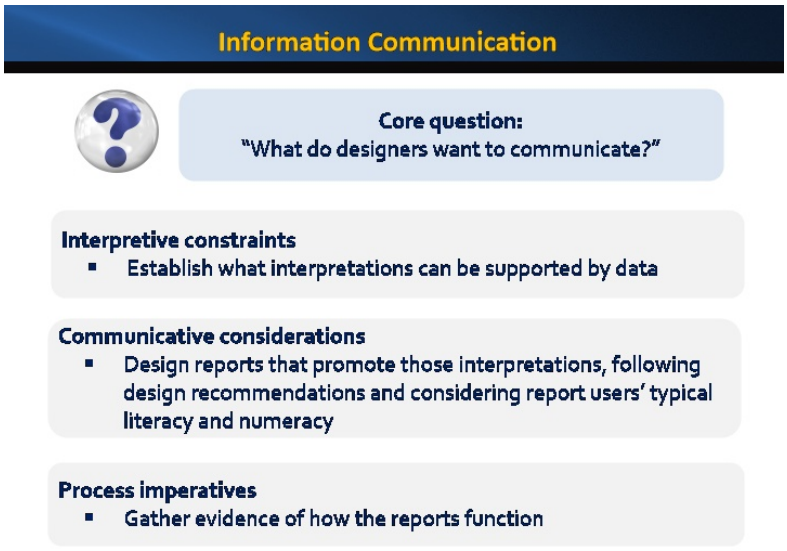
6.10 Conceptual layers of reporting



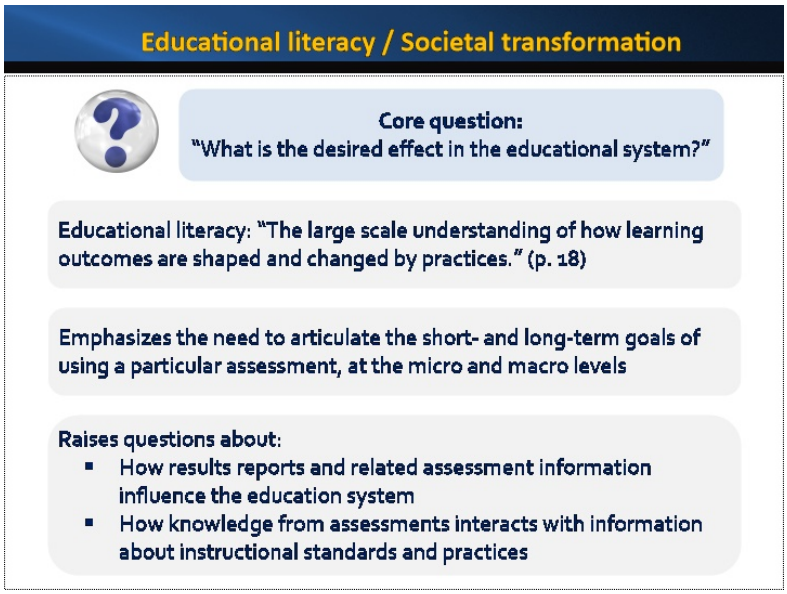
6.11 Conceptual Layers of Reporting



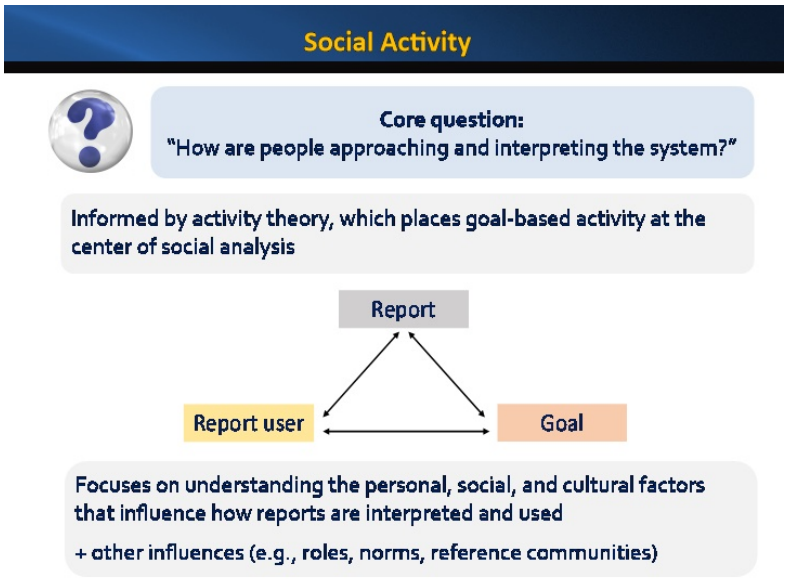
Information communication (Slide Layer)



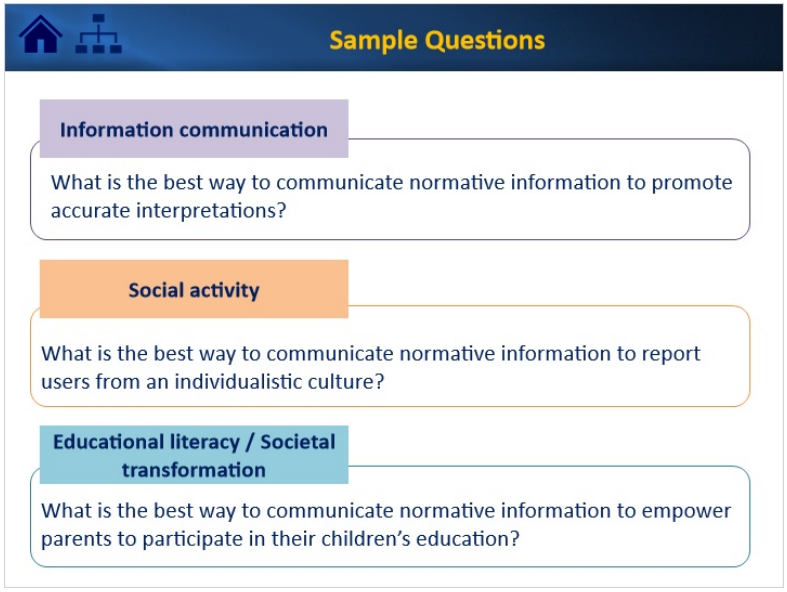
Educational literacy / Societal transformation (Slide Layer)



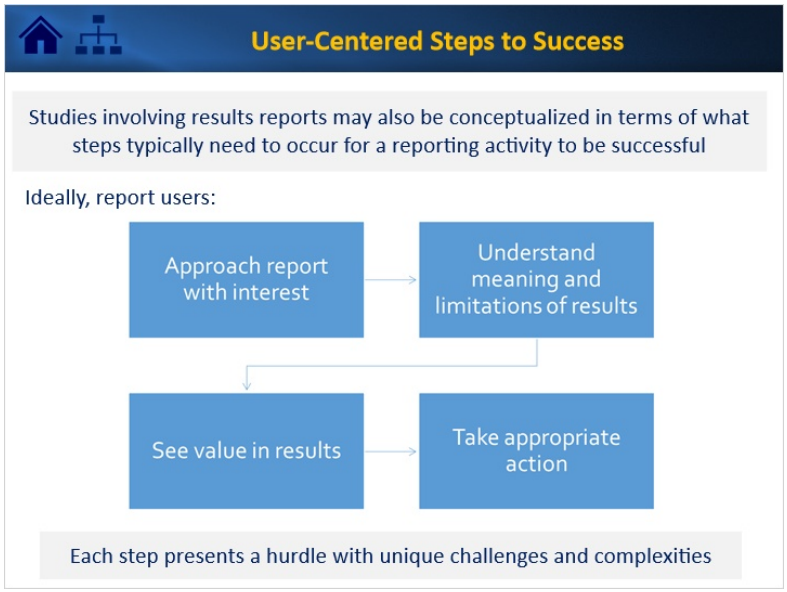
Social activity (Slide Layer)



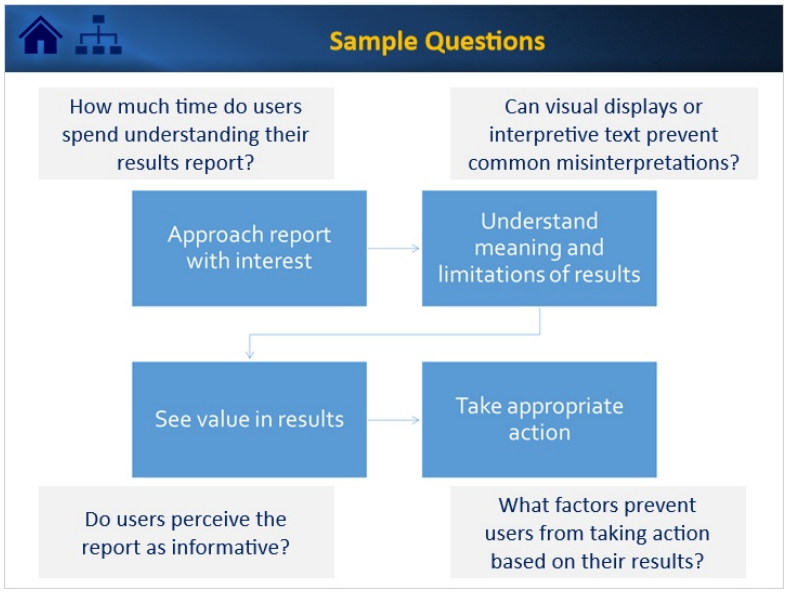
6.12 Sample Questions



6.13 User-Centered Steps to Success



6.14 Sample Questions



6.15 Sample Questions

Conclusion

Research on results reporting can evolve in multiple directions

For conceptual frameworks and ideas, consider:

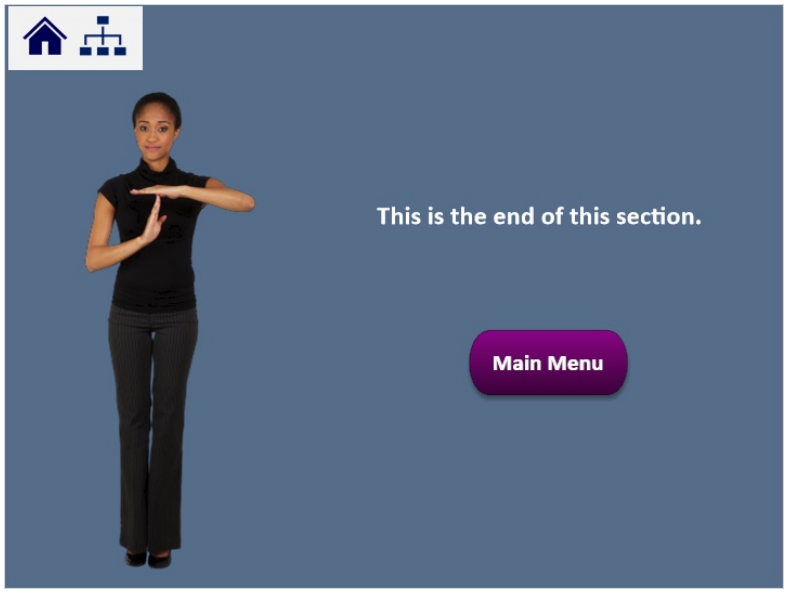
- Applications of communication theory traditions to results reporting
  - Gotch and Roduta Robert's (2018) review is a valuable starting point
- Behrens, DiCerbo, Murphy, and Robinson's (2013) Conceptual Layers of Reporting
- User-Centered Steps to Success

6.16 Personal Reflection #7

Learning Check: Personal Reflection #7

Choose a conceptual framework listed on slide 6.17, and briefly describe the major points of that framework as they relate to research on reporting that you might do or be interested in.

6.17 End of Section 5



6.18 Module Cover (END)