

# Annual Quality Control Report for

## Work Performed at the Center for Applied Linguistics

### Results from the ACCESS for ELLs 2.0 Quality Control Review in November 2017

January 2018



WIDA Assessment Team

Report prepared by the WIDA Consortium  
Wisconsin Center for Education Research  
1025 W. Johnson Street  
University of Wisconsin-Madison  
Madison, WI 53706

## ACCESS for ELLs 2.0 Quality Control Meeting

### Test Development and Psychometric Documentation Review

#### Center for Applied Linguistics

#### **Executive Summary**

The Center for Applied Linguistics (CAL) is WIDA's test development partner on the ACCESS for ELLs 2.0 project. Each year representatives from state education agencies within the WIDA consortium are invited to participate in a quality control (QC) meeting where CAL's test development processes and procedures are evaluated.

The SEA reviewers are provided with an overview of the ACCESS test development cycle and an introduction to the QC matrix on the first morning of the QC meeting. Following this orientation to the process, the SEA reviewers are also presented with an extensive set of documentation on all aspects of the test development process. This is broken down into the following categories for evaluation:

- Item and Test Development Documents
- Item Development Processes
- Training Materials Development
- Test Construction Procedures
- Test Proofs and Proof Handoff Process
- Item/Test Development System
- Psychometric Analysis and Research

For each of these categories, the SEAs independently review the documentation they are provided with. Each item is evaluated using one of the following ratings:

- Sufficient relevant evidence was presented
- Some relevant evidence was presented
- Little relevant evidence was presented
- No relevant evidence was presented

Overall, the evaluations of CAL's test development process from the SEA reviewers were very positive. The majority of categories and items evaluated were awarded the highest possible rating.

The two categories that received less favorable ratings were *test construction procedures* and *psychometric analysis and research*. As a result of these ratings and comments from reviewers, WIDA and CAL will address the following issues as priorities for improvement.

- Security protocols to ensure the secure transmission of all test content.
- Enhanced documentation that describes the creation of test maps.
- Enhanced documentation that describes which content on the test is new and which is continuing from previous administrations.
- Enhanced documentation that explains how test forms are assembled.
- More thorough account of who participates in the post-field test review and item selection meetings and how they are trained.
- A more detailed description of item pre-equating procedures.
- Enhanced documentation that describes how the annual technical report is created.

WIDA would like to express our appreciation to CAL for hosting the QC meeting and to the reviewers who gave up their time and expert feedback throughout the review. Full details of the methods and results from the meeting are contained within this report.

## **1. Background**

The Center for Applied Linguistics (CAL) has been WIDA's test development partner on the ACCESS for ELLs 2.0 (hereafter ACCESS) project since 2003. In addition to developing test content for ACCESS, CAL also provides psychometric and research services. WIDA and CAL collaborate annually on the development of new test content for ACCESS, analyses of field test content, item selection for new operational content, and a range of psychometric analyses. In addition, CAL creates rater training material for the ACCESS Speaking and Writing domain tests and authors the ACCESS Annual Technical Report.

## **2. Purpose**

The purpose of this quality control (QC) meeting is to review the processes and procedures that underlie CAL's test development and psychometric work on the ACCESS program. Representatives from WIDA and four state representatives from WIDA Consortium states participate in a review of the documentation that describes in detail CAL's test development and psychometric work. The state representatives are asked to evaluate the documentation using a checklist and rubric (hereafter QC matrix). The results of the evaluation are presented in this report and serve as evidence of the robustness of the processes and procedures that support the development and scoring of ACCESS. This QC review meeting is intended to serve as a thorough evaluation of industry standard test development procedures, from the creation of item specifications through to the publication of the final Annual Technical Report.

## **3. Scope**

On October 31 and November 1, 2017 a quality control group met at CAL in Washington DC. This group consisted of:

Kristine David, Indiana Department of Education  
Eugenia Krimmel, Pennsylvania Department of Education  
Jan Reyes, Georgia Department of Education  
Vince Verges, Florida Department of Education

WIDA was represented by:

Mark Chapman, Director of Test Development  
Carsten Wilmes, Director of Assessment Operations

The WIDA representatives facilitate the meeting but do not participate in the evaluation of CAL's processes and procedures. The four state education agency representatives (SEAs) complete the QC matrix, providing their independent evaluation of the test development and psychometric processes that underlie the ACCESS test program.

The rubric used by the SEAs to make their evaluations of the processes and procedures are provided in Appendix A.

## **4. Method**

The SEA reviewers are provided with an overview of the ACCESS test development cycle and an introduction to the QC matrix on the first morning of the QC meeting. Following this orientation to the process, the SEA reviewers are also presented with an extensive set of documentation on all aspects of the test development process. This is broken down into the following categories for evaluation:

- Item and Test Development Documents

This category includes key test development documents, such as test and item specifications, style guides, and test development plans. Also included within this category is a review of the qualifications held by ACCESS item writers.

- Item Development Processes

This category includes the processes by which writers submit test items, how these items are reviewed (both internally and at stakeholder review meetings), and how security is maintained.

- Training Materials Development

This category includes the processes and procedures that support the review of training materials to ensure they are consistent with current ACCESS test forms.

- Test Construction Procedures

This category includes procedures that support item storage (item banking), item selection, test map development, test form development and key checks.

- Test Proofs and Proof Handoff Process

This category includes the processes that support the creation and review of final test proofs.

- Item/Test Development System

This category includes a review of the organizational structure of the development team who create ACCESS and the security protocols that these staff adhere to.

- Psychometric Analysis and Research

This category includes the processes that support the analyses and selection of field test items, and the procedures that are followed during verification study and pre-equating, along with the production of the ACCESS annual technical report.

For each of these categories, the SEAs independently review the documentation they are provided with. Some topics are supported by presentations from relevant CAL staff who present

additional context to help the reviewers understand the documentation. Working in pairs, the SEAs complete their evaluations of the documentation using the QC matrix and each pair record a single rating for each item within the QC matrix. In addition, reviewers leave qualitative comments within the QC matrix where necessary, to explain their ratings. Each item is evaluated using one of the following ratings:

- Sufficient relevant evidence was presented
- Some relevant evidence was presented
- Little relevant evidence was presented
- No relevant evidence was presented

At the close of Day 1, and then again when all evaluations have been completed, the SEA reviewers and WIDA participants debrief on the main issues raised by the reviewers. At the conclusion of the QC meeting, the SEAs and WIDA staff briefly appraise CAL staff of the main findings.

The following section presents the results of the evaluations. Quantitative ratings and qualitative comments from all SEA reviewers are reported.

## **5. Results**

The results of the SEA evaluations are presented below. For each category, the evaluations of each pair of reviewers are tabulated and presented. Thus, two tables are presented within each category, first for Group A (two reviewers) and then Group B (two reviewers).

These tables are followed by the qualitative comments recorded by the SEA reviewers. These are presented in a consistent format, first showing the QC matrix descriptor that is being commented on, followed by the SEA comment in italics.

Qualitative comments from the reviewers that addressed wording of descriptors within the QC matrix are not included within this report. The qualitative comments reported below focus on issues identified with specific test development and/or psychometric processes and procedures.

### **5.1 Item and Test Development Documents**

As seen in Table 1, the quantitative ratings show that 100% of the evidence in this category was evaluated as sufficient by Group A.

Table 1. Summary Results for Item & Test Development Documents: Group A

	Sufficient relevant evidence was presented	Some relevant evidence was presented	Little relevant evidence was presented	No relevant evidence was presented
<b>Item &amp; Test Development Documents</b>	<b>100%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>
Content standards associated with blueprints & specs	100%	0%	0%	0%
Item specifications	100%	0%	0%	0%
Test specifications	100%	0%	0%	0%
Style guides	100%	0%	0%	0%
Maintenance & updating of development documents	100%	0%	0%	0%
Item/Test development plan	100%	0%	0%	0%
Test writer qualifications	100%	0%	0%	0%
Permissions Plan	100%	0%	0%	0%
Items, passage, and art approval process	100%	0%	0%	0%

The reviewers' qualitative comments are presented below.

- a) Content standards associated with blueprints & specs -- There is written documentation that associates test specifications with WIDA standards in each skill domain.  
*It would be helpful to have the standards attached. The documentation that associates the test with the standards is present, but I would have liked to see a copy of the standards attached as well.*
- b) Item specifications -- There is evidence that Item specifications are available and used by item writers.  
*Email show good feedback to item writers*
- c) Item specifications -- Test development staff have access and use items specifications in developing and reviewing new items.  
*Email show good feedback to item writers*

As seen in Table 2, the quantitative ratings show that 96% of the evidence in this category was evaluated as sufficient by Group B.

Table 2. Summary Results for Item & Test Development Documents: Group B

	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Item &amp; Test Development Documents</b>	<b>96%</b>	<b>4%</b>	<b>0%</b>	<b>0%</b>
Content standards associated with blueprints & specs	100%	0%	0%	0%
Item specifications	100%	0%	0%	0%
Test specifications	67%	33%	0%	0%
Style guides	100%	0%	0%	0%
Maintenance & updating of development documents	100%	0%	0%	0%
Item/Test development plan	100%	0%	0%	0%
Test writer qualifications	100%	0%	0%	0%
Permissions Plan	100%	0%	0%	0%
Items, passage, and art approval process	100%	0%	0%	0%

Qualitative Comments from Reviewers:

- a) Maintenance & updating of development documents-- There is an annual review within the test development staff of new forms of ACCESS to assure proper WIDA standard and ACCESS alignment.  
*There was a test development and review cycle in general; this evidence reflected that the review did happen.*

The quantitative ratings and comments from the reviewers indicate that the processes and procedures in place for ACCESS item development are very robust and provide assurance that the test is developed to a high standard of quality.

**5.2 Item Development Processes**

As seen in Table 3, the quantitative ratings show that 94% of the evidence in this category was evaluated as sufficient by Group A.



Table 3. Summary Results for Item Development Processes: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Item Development</b>	<b>94%</b>	<b>6%</b>	<b>0%</b>	<b>0%</b>
Edit/Re-write process	100%	0%	0%	0%
Internal item review process	100%	0%	0%	0%
Bias and Sensitivity review process	100%	0%	0%	0%
Committee review	80%	20%	0%	0%
Security and Intellectual Property	100%	0%	0%	0%

Qualitative Comments from Reviewers:

a) There is a process for ensuring that all copies of review materials are handled appropriately.

*How are items transmitted? Is the email secure?*

As seen in Table 4, the quantitative ratings show that 83% of the evidence in this category was evaluated as sufficient by Group B.

Table 4. Summary Results for Item Development Processes: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Item Development</b>	<b>83%</b>	<b>17%</b>	<b>0%</b>	<b>0%</b>
Edit/Re-write process	75%	25%	0%	0%
Internal item review process	100%	0%	0%	0%
Bias and Sensitivity review process	100%	0%	0%	0%
Committee review	60%	40%	0%	0%
Security and Intellectual Property	100%	0%	0%	0%

Qualitative Comments from Reviewers:

- a) Edit/Re-write process -- There is an item writer retraining and/or termination process based on the quality of submitted items.  
*The evidence provided seems to be an excerpt from a larger document. References are made to other sections (Section 2 and Section 6) that are not included in the provided evidence.*
- b) Internal item and passage review process -- There is an internal procedure, consistent with test and item specifications, that assures item and passage appropriateness.  
*Very well developed process*
- c) External Content Committee review -- There is a system in place for training Content Review Committee facilitators, monitoring their performance, and providing feedback.  
*Evidence is provided related to training facilitators, but not for monitoring their performance and providing feedback. Most of the evidence is related to the meeting itself (agenda, checklists, etc.)*

The quantitative ratings are again generally strong within this category; however, the ratings indicate that committee review processes and procedures were not evaluated as positively as the other items on the checklist. The single comment on this activity indicates that the reviewers wanted to see more evidence regarding the performance of the committee facilitators and the feedback provided to them.

### 5.3 Training Materials Development

As seen in Table 5, the quantitative ratings show that 100% of the evidence in this category was evaluated as sufficient by Group A.

Table 5. Summary Results for Training Materials Development: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Training Materials Development</b>	<b>100%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>
Administration training process and procedures	100%	0%	0%	0%

Qualitative Comments from Reviewers:

- a) Administration training process and procedures -- Procedures are in place to review test administration training materials to assure they are consistent with current forms of ACCESS.

*One part of the Phase 6 narrative states that CAL creates much of the content, but the table and this criteria seems to indicate CAL's role is primarily that of reviewer.*

As seen in Table 6, the quantitative ratings show that 100% of the evidence in this category was evaluated as sufficient by Group B.

Table 6. Summary Results for Training Materials Development: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Training Materials Development</b>	<b>100%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>
Administration training process and procedures	100%	0%	0%	0%

There were no pertinent qualitative comments from the Group B reviewers in this category.

The quantitative ratings from the reviewers indicate that there are no issues with the processes and procedures that support CAL's work on the ACCESS training materials.

#### **5.4 Test Construction Procedures**

As seen in Table 7, the quantitative ratings show that 71% of the evidence in this category was evaluated as sufficient by Group A.

Table 7. Summary Results for Test Construction Procedures: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Test Construction</b>	<b>71%</b>	<b>14%</b>	<b>5%</b>	<b>0%</b>
Item banking solution	67%	33%	0%	0%
Field test item selection & implementation process	67%	0%	0%	0%
Test map development & approval process	33%	33%	33%	0%
Test directions and ancillary test material development process	67%	0%	0%	0%
Test form review process	83%	17%	0%	0%
Test key development and backcheck process	100%	0%	0%	0%

Group A reviewers did not respond to two items in this category. Items D6 and D12 were not completed on the QC Matrix, therefore two of the rows in Table 7 do not total 100% (Field test item selection & implementation process; Test directions and ancillary test material development process)

Qualitative Comments from Reviewers:

- a) Item banking solution -- The ACCESS item storing system has secure storage of items and item information with checks in place to assure only relevant parties have access. A named party is responsible for maintaining test security.  
*I was expecting to see more about the security protocols. For item reviews, are emails sent securely? Appendix B of Test Security Guide should be cited.*
- b) Test map development & approval process -- A process is in place for creating test maps for each ACCESS form.  
*I don't see the documentation or the process on how you get to this point.*
- c) Test map development & approval process -- ACCESS Test maps are checked with test specifications and WIDA standards to assure appropriate alignment.  
*Not clear enough how operational and FT interact. Which OP items stay, and which are retired?*
- d) Test form review process -- A process is in place to take all revised field tested items and existing ACCESS operational items, as well as associated ancillary documents, and put them together in final form.  
*This doesn't specifically state how these forms are put together.*

As seen in Table 8, the quantitative ratings show that 81% of the evidence in this category was evaluated as sufficient by Group B.

Table 8. Summary Results for Test Construction Procedures: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Test Construction</b>	<b>81%</b>	<b>19%</b>	<b>0%</b>	<b>0%</b>
Item banking solution	100%	0%	0%	0%
Field test item selection & implementation process	100%	0%	0%	0%
Test map development & approval process	33%	67%	0%	0%
Test directions and ancillary test material development process	100%	0%	0%	0%
Test form review process	67%	33%	0%	0%
Test key development and backcheck process	100%	0%	0%	0%

There were no pertinent qualitative comments from Group B reviewers in this category.

The reviewers, particularly in Group A, reported some reservations about the processes and procedures that support ACCESS test construction. Specifically, concern was expressed about security protocols for the transmission of item reviews that may contain test content, documentation that describes the creation of test maps, continuation and renewal of operational test content, and the assembly of test forms.

This feedback has been shared with CAL and improvements in the documentation that account for these processes described above will be a priority for 2018.

### 5.5 Test Proofs and Proofs Handoff Process

As seen in Table 9, the quantitative ratings show that 80% of the evidence in this category was evaluated as sufficient by Group A.

Table 9. Summary Results for Review of Test Proofs & Proofs Handoff Process: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Review of Test Proofs &amp; Proofs Handoff Process</b>	<b>80%</b>	<b>20%</b>	<b>0%</b>	<b>0%</b>

Qualitative Comments from Reviewers:

- a) Review of Test Proofs & Proofs Handoff Process -- A secure process is in place to send final proofs to WIDA for final approval and sign off.  
*How is this a secure process? In other words, what is the delivery method?*
- b) Review of Test Proofs & Proofs Handoff Process -- A secure file transfer and recording protocol is in place to send final proofs to the production contractor.  
*From pg. 16 of the Security Guide, as it relates to item development: "The item writer does not send any work by return email, but rather submits all work via the password-protected Online Item Writer database." This policy is apparently being routinely violated with respect to edits, so either remove this from the guide, modify it, or adhere to the policy as it was written.*

As seen in Table 10, the quantitative ratings show that 80% of the evidence in this category was evaluated as sufficient by Group B.

Table 10. Summary Results for Review of Test Proofs & Proofs Handoff Process: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Review of Test Proofs &amp; Proofs Handoff Process</b>	<b>80%</b>	<b>20%</b>	<b>0%</b>	<b>0%</b>

Qualitative Comments from Reviewers:

- a) Review of Test Proofs & Proofs Handoff Process -- A secure file transfer and recording protocol is in place to send final proofs to the production contractor.  
*Who prints the paper forms? I don't see any reference to printing/transferring files to the printed.*

Reviewers awarded generally positive ratings in this category, although reviewers expressed some minor reservations about how CAL securely transfers test content and test materials.

**5.6 Item/Test Development System**

As seen in Table 11, the quantitative ratings show that 90% of the evidence in this category was evaluated as sufficient by Group A.

Table 11. Summary Results for Item/Test Development System: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Item/Test Development System</b>	<b>90%</b>	<b>10%</b>	<b>0%</b>	<b>0%</b>
Organizational structure	83%	17%	0%	0%
Staff qualifications and experience	100%	0%	0%	0%
Security	100%	0%	0%	0%

Qualitative Comments from Reviewers:

- a) Organizational structure -- There is a mechanism in place to monitor the effectiveness of test development staff.  
*You may want to include a copy of the performance evaluation rubric utilized by CAL for internal test development staff.*

As seen in Table 12, the quantitative ratings show that 100% of the evidence in this category was evaluated as sufficient by Group B.

Table 12. Summary Results for Item/Test Development System: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Item/Test Development System</b>	<b>100%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>
Organizational structure	100%	0%	0%	0%
Staff qualifications and experience	100%	0%	0%	0%
Security	100%	0%	0%	0%

Qualitative Comments from Reviewers:

- a) Organizational structure -- There is a clear, articulated structure (schematic) for the item/test development process.  
*Very well articulated.*

Reviewers did not express any concerns about the processes and procedures described in documentation within this category. One minor addition to the documentation was requested but the processes and procedures themselves were evaluated highly.

## 5.7 Psychometric Analysis and Research

As seen in Table 13, the quantitative ratings show that 47% of the evidence in this category was evaluated as sufficient by Group A.

Table 13. Summary Results for Psychometric Analysis and Research: Group A

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Psychometric Analysis and Research</b>	<b>47%</b>	<b>53%</b>	<b>0%</b>	<b>0%</b>
Annual Field Test Cycle	67%	33%	0%	0%
Post Field Test Review and Operational Item Selection	0%	100%	0%	0%
Spring Verification Study Procedures	67%	33%	0%	0%
Pre-Equating Procedures	33%	67%	0%	0%
Annual Tech Report Development	67%	33%	0%	0%

Qualitative Comments from Reviewers:

- a) Annual Field Test Cycle -- A clear line of sign-offs and procedures for approval are documented and used in the implementation.  
*Sign-offs are provided, but the process is not within these artifacts.*
- b) Post Field Test Review and Operational Item Selection -- A system is in place to ensure all steps in the post field test data review and operational item selection process are completed.  
*Not sure what the "rehearsal" meeting is, or how it fits into the process.*
- c) Post Field Test Review and Operational Item Selection -- Quality control procedures are in place to ensure proper implementation.  
*It is not clear who the participants are in this review. If these are educators, do they engage in some sort of training prior to rendering their judgments?*
- d) Spring Verification Study Procedures -- A clear line of sign-offs and procedures for approval are documented and used in the implementation.  
*No procedures included- only the actual sign-offs.*
- e) Pre-Equating Procedures -- A system is in place to ensure all steps in the pre-equating procedures are completed.  
*It isn't clear how the pre-equating is done for paper forms.*



As seen in Table 14, the quantitative ratings show that 73% of the evidence in this category was evaluated as sufficient by Group B.

Table 14. Summary Results for Psychometric Analysis and Research: Group B

<b>WIDA Test Development Process Checklist</b>	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
<b>Psychometric Analysis and Research</b>	<b>73%</b>	<b>27%</b>	<b>0%</b>	<b>0%</b>
Annual Field Test Cycle	100%	0%	0%	0%
Post Field Test Review and Operational Item Selection	33%	67%	0%	0%
Spring Verification Study Procedures	100%	0%	0%	0%
Pre-Equating Procedures	67%	33%	0%	0%
Annual Tech Report Development	67%	33%	0%	0%

Qualitative Comments from Reviewers:

- a) Post Field Test Review and Operational Item Selection -- A clear line of sign-offs and procedures for approval are documented and used in the implementation.  
*Would the Decision Recorder and Decision Observer not be part of the line of approval?*
- b) Pre-Equating Procedures -- A system is in place to ensure all steps in the pre-equating procedures are completed  
*A description file would have been helpful here, similar to the other phases.*
- c) Pre-Equating Procedures -- A clear line of sign-offs and procedures for approval are documented and used in the implementation.  
*Helpful to have an actual procedure, not just emails from individuals*
- d) Annual Tech Report Development -- Quality control procedures are in place.  
*Quality control procedures are described for the data, but not for the development of the report itself.*

The reviewers awarded lower ratings in this category (psychometric analysis and research) than in the other categories covered in this QC meeting. Both groups expressed a hope for better documentation of the field test review and item selection processes. The documentation that accounts for pre-equating procedures was also seen as less than sufficient, as was the case for the development of the annual technical report.

This feedback has been shared with CAL and improvements in the documentation that account for these processes described above will be a priority for 2018.

## **6. Summary & Conclusion**

The majority of the processes and procedures reviewed by the SEAs received positive ratings. The following categories received positive evaluations from both groups, with no significant concerns raised.

- Item and Test Development Documents
- Item Development Processes
- Training Materials Development
- Test Proofs and Proof Handoff Process
- Item/Test Development System

Reviewers were more critical of the following two categories.

- Test Construction Procedures

Aspects of test construction that reviewers expressed reservations about included security protocols for the transmission of item content, documentation that describes the creation of test maps, documentation that explains test content that is newly operational and content that is continuing, and details of how the test forms are assembled.

- Psychometric Analysis and Research

Documentation of processes and procedures in this category were not seen as sufficient by the reviewers. In particular, the field test review and item selection processes, the pre-equating procedures, the development of the annual technical report were reviewed critically.

Overall, WIDA views the findings of the 2017 CAL QC meeting as providing broad support for the claim that ACCESS is developed according to rigorous processes and procedures. The highly detailed documentation provided to the reviewers and scrutinized for two days was evaluated as sufficient in the vast majority of cases. These broadly positive evaluations serve as compelling evidence for the thoroughness of the work that support the creation of ACCESS for ELLs 2.0. The relatively small number of areas where the documentation was seen as insufficient and the processes and procedures that support ACCESS were questioned will be addressed directly between WIDA and CAL.

## **7. Recommendations**

Based on the evaluation presented above, WIDA and CAL will address the following issues as priorities for improvement.

- Security protocols to ensure the secure transmission of all test content.
- Enhanced documentation that describes the creation of test maps.
- Enhanced documentation that describes which content on the test is new and which is continuing from previous administrations.
- Enhanced documentation that explains how test forms are assembled.
- More thorough account of who participates in the post-field test review and item selection meetings and how they are trained.
- A more detailed description of item pre-equating procedures.
- Enhanced documentation that describes how the annual technical report is created.

Appendix A: Checklist Rubric

<b>WIDA Test Development Checklist Rubric</b>				
<b>Indicator 1: Products</b>				
	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
The observed products, materials and/or artifacts are present, understandable and function to support their purpose.				
<b>Indicator 2: Process, Procedures and Protocols</b>				
	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
The observed processes, procedures or protocols are present, understandable, executable, and support the purpose for which they were developed.				
<b>Indicator 3: Personnel</b>				
	<b>Sufficient relevant evidence was presented</b>	<b>Some relevant evidence was presented</b>	<b>Little relevant evidence was presented</b>	<b>No relevant evidence was presented</b>
Person assigned and observed qualifications and/or interviews of assigned persons are consistent with their assignment(s) and reflect the capacity to accomplish their assigned tasks.				
	<b>NOTE:</b> If items' personnel, artifacts or processes in your judgment are acceptable but no formal written documentation exists, list those items as " <i>Some relevant evidence was presented</i> " and annotate the lack of documentation on the comment field.			