

## Executive Summary of the CCCAssess Review

The CCC Assessment Committee received the application of the CCCAssess submitted by the Common Assessment Initiative (CAI) on July 30, 2016. The *Standards, Policies and Procedures for the Evaluation of Assessment Instruments Used in the California Community Colleges* (4th Ed., 2001; hereafter referred to as *CCC Assessment Standards*) was used to guide the Assessment Committee's review of the CCCAssess application. Specifically, the test standards delineated in the *CCC Assessment Standards* that are applicable to a second-party test were used to evaluate the CCCAssess. The Assessment Committee met on August 10 and 11, 2016, to discuss the application. A consensus of non-approval recommendation for the CCCAssess in its current state was made by the Assessment Committee at the conclusion of the review meeting.

This executive summary briefly summarizes the key findings from that review. We encourage the readers to consult the full report titled *Review of the CCCAssess English Language Arts and Mathematics Assessments* (CCC Assessment Committee, August 2016) for a detailed version of review comments and suggestions.<sup>i</sup>

### Establishing Intended Uses and Interpretations for CCCAssess

- Descriptions of the intended uses and interpretations of results for the CCCAssess instruments (e.g., test score or ability measure) must be clearly specified.
- Evidence in support of the intended usage (e.g., scaling evaluation and dimensionality<sup>ii</sup>) must be provided.

### Content-Related Validity Evidence

- Details and results of a properly conducted alignment study must be provided.
- The content coverage summary for the items in the operational item bank (i.e., items found to be aligned, free of bias, and functioning as expected in the pilot) must be provided for all subject areas.
- The content coverage of the operational items is inadequate for some competency categories and/or difficulty levels. More items may be needed for the operational item bank; otherwise, CAI must be explicit about the content limitation of the tests.
- Subjectively judged item difficulty levels should be empirically verified.
- A representative sample of items must be available upon request by potential test users (local colleges).
- The computer adaptive algorithm (starting point, item/testlet selection, termination, test scoring, and item exposure control, etc.) must be fully described, and its rationale and supporting evidence must be provided. What constraints (content coverage and sufficiency, standard error of measurement, and number of testlets administered, etc.) are implemented with the adaptive algorithm and how content coverage and sufficiency are operationalized in the adaptive algorithm should be well defined for test users.
- The branching of English and ESL testing and their testing progression must be clearly articulated.
- The content of the education background survey and a description for how responses are synthesized must be provided. There might be legal and fairness concerns for asking some of the demographic questions. The evidence for the efficacy of using the education background survey for determining the entry point must be provided.
- A plan for the items in the operational item bank to be periodically reviewed should be provided.

### **Fairness Evidence**

- A logical bias panel review must be properly conducted and documented. The follow-up investigations and actions toward flagged items must be described.
- A detailed summary for the empirical differential item functioning (DIF) analysis must be provided. Additional attention should be paid to sample sizes of the subgroups included in the analysis.
- All items in the operational item bank must have undergone logical bias review, and at least 80% of the items in the operational item bank must have appropriate DIF analyses completed. Test developers must address any bias concerns arising out of logical review or empirical analyses before the items can be considered for use in the operational item bank.

### **Criterion-Related or Consequential-Related Validity Evidence**

- Validity evidence must be collected for the operational version of the assessments.
- Details for the participants, courses surveyed, and data collection procedures must be provided for users to evaluate the relevancy of the validity evidence.
- The minimally acceptable correlation for supporting criterion-related validity is .35 (*CCC Assessment Standards*, p.19).
- Construct-irrelevant barriers to students' test performance on CCCAssess should be investigated and addressed.

### **Reliability**

- Reliability evidence must be collected for the operational version of the assessments.
- Standard errors of measurement must be provided.
- Item types that resulted in polytomous scores should be described. If item scoring involved rubrics and/or machine scoring with artificial intelligence, additional reliability and/or validity evidence for scoring accuracy must be provided.
- Test-retest reliability must be provided in order to receive full approval.

### **Testing Special Groups**

- Commonly requested accommodations must be available (e.g., paper-and-pencil forms<sup>iii</sup> and Braille)
- There should be clear accommodation policies/guidelines for test users to follow.

---

<sup>i</sup> We note that the review report was the CCC Assessment Committee's report rather than Buros' report.

<sup>ii</sup> The dimensionality issue emerged during the August 23 follow-up meeting with CAI in Sacramento.

<sup>iii</sup> A separate assessment review application must be submitted for the approval of the paper-and-pencil forms. The same standards (e.g., fairness, content validity, consequential-/criterion-related validity, and reliability) must be addressed for the paper-and-pencil forms.