OFFICE OF CHIEF ACCOUNTABILITY OFFICER
Summary of State Board of Education Agenda Items
May 1, 2024

OFFICE OF STUDENT ASSESSMENT

01. Action: Approval of the recommended cut scores on the Cambridge International General Certificate of Secondary Education (IGCSE) as end-of-course assessments for English, Mathematics, Biology, and History [Goals 1 and 2 – MBE Strategic Plan]

Background Information: In 2022 the United States Department of Education approved Cambridge IGCSE assessments in First Language English, Mathematics, and Biology for use as Locally Selected, Nationally Recognized High School Assessments (LSNRHSA) in Corinth School District in Corinth, Mississippi. As part of the process of implementing these three assessments, and additionally the Cambridge IGCSE assessment in American History, in place of the statewide end-of-course tests for English II, Algebra I, Biology, and US History, Cambridge has designed a robust process that allows for Cambridge IGCSE scores in these content areas to be translated into Mississippi end-of-course performance levels.

The first stage of the process was a statistical linking study using test data from Corinth School District, where students had taken both Cambridge IGCSE assessments and MAAP assessments in two (2) separate academic years. The 2022 linking study, using 2020 – 2021 data and replicated with 2018 – 2019 pre-pandemic data, established strong content alignment between the Cambridge IGCSE assessments and the Mississippi College- and Career-Readiness Standards (CCRS).

The second stage consisted of a standards validation exercise whose objective was to propose defensible cut scores at the four (4) performance level thresholds in each content area. This review was based on scrutiny of candidate performance from the June 2023 administration, considering test items and student work in relation to established performance level descriptors. The standards validation workshop took place on March 26 – 28, 2024, at Corinth Elementary School in Corinth, Mississippi, and involved a committee of 28 panelists across the four (4) content areas.

As Cambridge IGCSE assessments include a significant number of constructed response and multi-part items and have a strong focus on evaluating candidate-generated evidence, a novel methodology was developed for the exercise, drawing on both Body of Work (BoW) and Bookmark approaches to standard setting Panelists for English and History followed a BoW method with three (3) distinct rounds. Panelists for Mathematics and Biology followed a blended approach, drawing on Bookmark and BoW methods with three (3) distinct rounds. The results demonstrate that these cut scores lead to broadly similar outcomes for the full cohort taking Cambridge IGCSE assessment compared with taking the MAAP.

This item references Goals 1 and 2 of the *Mississippi Board of Education Strategic Plan.*

Recommendations: Approval

Back-up material attached

Administration of Cambridge IGCSE as a LSNRHSA in Mississippi

Proposed cut scores for Cambridge IGCSE as end-of-course assessments for English, Mathematics, Biology and History

Stuart Gallagher and Dr. Justin Coy

April 2024

In 2022 the US Department of Education approved Cambridge IGCSE assessments in First Language English, Mathematics and Biology for use as Locally Selected, Nationally Recognized High School Assessments (LSNRHSA) in Corinth School District, Mississippi. As part of the process of implementing these three assessments, and additionally the Cambridge IGCSE assessment in American History[1], in place of the statewide end-of-course tests for English II, Algebra I, Biology, and US History, Cambridge has designed a robust process that allows for Cambridge IGCSE scores in these content areas to be translated into Mississippi end-of-course performance levels.

The first stage of the process was a statistical linking study using test data from Corinth School District, where students had taken both Cambridge IGCSE assessments and MAAP tests in two separate academic years. The 2022 linking study, using 2020/21 data and replicated with 2018/19 pre-pandemic data, established strong content alignment between the Cambridge IGCSE assessments and the Mississippi CCRS. This study provided suggested ranges within which appropriate cut scores would likely sit.

The second stage was a standards validation exercise whose objective was to propose defensible cut scores at the four performance level thresholds in each content area. This review was based on scrutiny of candidate performance from the June 2023 administration, considering test items and student work in relation to established performance level descriptors.

The standards validation workshop took place on March 26-28, 2024, at Corinth Elementary School in Corinth, Mississippi and involved a committee of 28 panelists across the four content areas. As Cambridge IGCSE assessments include a significant number of constructed response and multi-part items and have a strong focus on evaluating candidate-generated evidence, a novel methodology was developed for the exercise, drawing on both Body of Work (BoW) and Bookmark approaches to standard setting. Furthermore, and in line with the aim of ***validating*** rather than ***setting*** standards, this method treated the linking study as the range-finding stage, therefore allowing the workshop to move directly to the pinpointing stage. Panelists for English and History followed a BoW method with three (3) distinct

[1] This executive summary includes outcomes related to IGCSE American History. The submission of the full standards validation report will not include IGCSE American History as that was not submitted as part of the Locally Selected Nationally Recognized High School Assessment submission.

Table 1 presents the cut scores recommended by the panels in the four content areas, and Table 2 presents the impact data for the recommended cut scores, based on the June 2023 cohort in Corinth School District.

Content area	Basic	Passing	Proficient	Advanced
English	25	31	48	66
Mathematics	20	26	43	72
Biology	18	26	45	70
History	15	23	38	57

Table 1: Proposed cut scores at each performance level, expressed in Cambridge scores

Content area	N	Minimal	Basic	Passing	Proficient	Advanced
English	165	9.1%	1.8%	14.6%	32.7%	41.8%
	(cum.)	*(100%)*	*(90.9%)*	*(89.1%)*	*(74.5%)*	*(41.8%)*
Mathematics	148	37.2%	9.4%	25.0%	25.7%	2.7%
	(cum.)	*(100%)*	*(62.8%)*	*(53.4%)*	*(28.4%)*	*(2.7%)*
Biology	157	34.4%	24.8%	23.6%	14.7%	2.5%
	(cum.)	*(100%)*	*(65.6%)*	*(40.8%)*	*(17.2%)*	*(2.5%)*
History	151	11.3%	9.2%	32.5%	27.8%	19.2%
	(cum.)	*(100%)*	*(88.7%)*	*(79.5%)*	*(47.0%)*	*(19.2%)*

Table 2: Impact data for the recommended cut scores – percentage achieving each performance level and cumulative percentage achieving each performance level or above

Procedure

For all panelists, the process involved three distinct rounds of scrutiny and judgements, with discussion and feedback throughout. The outline structure of the process is presented in Figure 1.

Round	Objective	English and History	Mathematics and Biology
1	Confirm guardrails – proposed ranges	BoW	Bookmark
2	Identify acceptable cut score at each level		
3	Validate cut scores for varied performances		BoW

Figure 1: Outline of the two approaches taken for different content areas

Panelists for English and History followed a BoW method with three distinct rounds. The detail of these rounds is as follows:

- Round 1: Presented with a small number of student profiles spanning the proposed range of plausible cut scores. Panelists made a 'yes~no' determination that performance at the threshold between performance levels was in the range.
- Round 2: In response to the panel's aggregated round 1 judgements, panelists reviewed a single profile at each Cambridge score point, in order to identify the profile best representing threshold performance. Panelists worked individually, then discussed views in small groups, before making individual judgements that were aggregated and the range of views considered. Impact data was shared with panelists, including the proposed cut scores.
- Round 3: Having reached broad agreement on a single cut score for each performance level, panelists reviewed further profiles on the same Cambridge score point, and potentially one or two points either side, in order to confirm the cut score and reach group consensus.

Panelists for Mathematics and Biology followed a blended approach, drawing on Bookmark and BoW methods, again with three distinct rounds. The detail of these rounds is as follows:

- Round 1: Using the ordered item booklet (OIB), panelists reviewed a specified set of items representing the proposed range of plausible cut scores. They then made a 'yes~no' determination that performance at the threshold between performance levels was in the range.
- Round 2: In response to the panel's aggregated round 1 judgements, panelists reviewed the ranges of items in more detail, in order to identify the profile best representing threshold performance; this involved a number of mini-rounds. In each mini-round, panelists worked individually, then discussed views in small groups, before making individual judgements that were aggregated and the range of views considered. Impact data was shared with panelists at each stage, alongside the proposed cut score this included 'pseudo profiles', a cluster of items where a typical student would likely score half the points, which informed the next round of scrutiny and judgements.
- Round 3: Having reached broad agreement on a single cut score for each performance level, panelists then reviewed a student profile at the proposed Cambridge score point, and potentially one or two points either side, in order to confirm the cut score and reach group consensus.

Throughout the process, panelists responded positively to the impact data presented between operational rounds and used it judiciously when revisiting their previous judgements. Generally, the presentation of impact data led to panelists selecting items or cut scores that were slightly lower, though they were reminded that their judgements should primarily focus on performance standards.

For Mathematics and Biology, round 3 was particularly significant as it moved from reviewing student performance in the abstract (i.e., items presented in an OIB), to the concrete, illustrating how students had interacted with the assessment and what a full test performance at given Cambridge score point represented. As a result of this additional, more holistic evidence of student performance, the Biology panel recommended marginally lower cut scores for Proficient and Advanced.

As part of round 3, final impact data was again shared with panelists and views collected through evaluation forms. All panelists endorsed the proposed cut scores from round 3, agreeing that they accurately reflected the collective expert judgement of student test performance in relation to the state's performance level descriptors and experience of student performance more generally. However, some panelists voiced concerns about the impact of the proposed cut scores, specifically that percentages achieving each performance level or above were, in places, lower than they might ordinarily expect.

Overall, the proposed cut scores are within a reasonable range to those suggested by the statistical linking study, which was used to derive the proposed ranges of plausible scores. Table 3 presents the cut scores recommended by the panels in the four content areas alongside the cut scores suggested by statistical linking, with confidence intervals. The proposed cut scores are generally a little lower, except the Basic cut score, which is a little higher. The results demonstrate that these cut scores lead to broadly similar outcomes for the full cohort taking Cambridge IGCSE assessment compared with taking the MAAP.

Content area	**Basic**		**Passing**		**Proficient**		**Advanced**	
	Rec.	**Link.**	**Rec.**	**Link.**	**Rec.**	**Link.**	**Rec.**	**Link.**
English	25	23 *(20-26)*	31	35 *(31-39)*	48	50 *(46-54)*	66	68 *(65-71)*
Mathematics	20	20 *(17-23)*	26	29 *(26-32)*	43	43 *(40-46)*	72	70 *(67-73)*
Biology	18	17 *(14-20)*	26	27 *(24-30)*	45	46 *(43-53)*	70	76 *(71-81)*
History	15	14 *(9-19)*	23	27 *(24-30)*	38	41 *(38-44)*	57	66 *(61-71)*

Table 3: Proposed cut scores from standards validation (Rec.) compared to the cut scores suggested by statistical linking (Link.), with confidence intervals