APPENDIX Z

COGNITIVE SKILLS REVIEW AND UPDATE PROCESS

Appendix Z: Cognitive Skills Review and Update Process

Starting in 2020, the Department of Elementary and Secondary Education (DESE) test developers, as well as the contractor's test developers, reviewed and revised guidance on writing items to varying levels of cognitive complexity, as well as the definitions of cognitive complexity levels themselves. Within the Massachusetts Comprehensive Assessment System (MCAS) program, cognitive complexity is operationalized in terms of cognitive "skill levels" that are derived from Norman Webb's depth of knowledge framework.

The goal of review and revision work for the ELA and math cognitive skills was to do the following:

- Develop a deeper and shared understanding of each cognitive skill level.
- Update the cognitive skills attribute handouts.
- Create cognitive skill target ranges for operational tests.
- Develop grade-level examples for each cognitive skill level using released MCAS items.
- Provide guidance and training to item developers and item review committees, including the Assessment Development Committees (ADCs), on the revised cognitive skill level materials to support all future development.

Cognitive Skills Review and Revision Process

In 2018 and 2019, researchers from Boston College conducted two alignment studies, one study for grades 3–8 and another for grade 10. These two studies suggested several areas of improvement, including improvements related to cognitive complexity. In response, a series of meetings were held to review and revise the approach to cognitive complexity taken with the MCAS program. These meetings were conducted separately for English Language Arts (ELA) and mathematics but had the same overall structure and steps. Each meeting was co-facilitated by DESE and Cognia test development leads. As explained below, these steps started with a review of the results of the alignment studies and ended with the revision of the cognitive skills attribute handout, creation of targets for each grade level test and the collation of grade-level example items for each cognitive skill.

In each content area, test development leads facilitated review of (a) the cognitive skill guidance document, (b) the complexity skill levels assigned through the alignment studies, (c) the complexity skill levels assigned to the items through the development process and (d) the items and associated stimulus materials used within the alignment studies. Note that the alignment studies examined the 2017 operational form for grades 3–8 and the 2019 operational form for grade 10. Key in this review was understanding differences between (b) and (c), the complexity skill levels assigned through the alignment studies and those assigned through the item development process. Each content area used slightly different processes, but ultimately arrived at the same set of materials to inform future development.

English Language Arts

For ELA, grade-level team members independently assigned cognitive skills to each item on each grade level test, based on the language within the pre-existing cognitive skills guidance documents. The levels assigned were then compared to those from the alignment study as well as those originally assigned to the items to facilitate conversation around differences in assignments and as well as the assignments themselves. These conversations were used to inform the revision of the cognitive skills attribute handout, as well as targets for cognitive skills on the operational form and training materials.

Three meetings were held by the test developers. In preparation for the first meeting, test developers read the passages and accompanying items that were included in part of the alignment study. For grades 3–8, there were three passage sets from the 2017 tests and for grade 10, there were four passage sets from the 2019 test. During the meeting, test developer leads met with grade level test developer teams to review the current language of the cognitive skills within the cognitive skills handout. The teams then reviewed a selection of released items and discussed differences between the originally assigned cognitive skill level and those from the alignment study. The teams then assigned cognitive skill levels to each item on the grade level test. Following this meeting, the test developer leads combined the cognitive skill level assignments from the grade level teams with the assignments from the original item development along with those from the alignment study.

In the second meeting, the grade level teams reviewed the combined cognitive skill levels assignments. In this review the teams were asked to consider the following questions:

- Why might there be a difference of cognitive level assignment between groups?
- Why was a cognitive level selected for an item?

The grade-level teams then selected items to share with the whole group. They shared how each item was assigned to a cognitive skill level and sought feedback on items where discrepancies occurred. As part of the discussion, the teams were also asked to think about attributes for each cognitive skill to prepare for the next meeting, which involved building consensus for the alignment study document. Following the second meeting, the test content leads determined the target distribution of score points by cognitive skill level based on careful examination of the standards, operational item banks and cognitive skills handout.

In the third and final meeting, all the ELA test developers joined a whole group conversation focused on revisions to the cognitive skill attributes handout, as well as the target distribution of score points by cognitive skill level for each operational test. As part of this discussion, test developers were asked to think about how the target distribution should be considered during item development and forms construction. The test developer leads also worked after the meeting to finalize the cognitive skill attributes handout and create training materials for ADCs on the cognitive skill attributes handout. These training materials included carefully curated exemplar items at each cognitive skill level.

Mathematics

Similarly, in mathematics a team reviewed the cognitive skill levels assigned through the alignment studies, as well as the original assessments from test development. The team discussed differences in cognitive skill assignments to again develop a shared understanding of the differences as well as the levels themselves. Notably, the team spent a great deal of time examining variations in understanding and assignment of cognitive skill categories 2 and 3. The team also engaged with staff at education departments in Connecticut, Kentucky, North Carolina, Iowa, and Minnesota to understand other states' approaches to cognitive complexity. The team also drew on Karin Hess' "A Guide for Using Webb's Depth of Knowledge with Common Core State Standards" (Hess, 2013). As with ELA, this work was used to update the cognitive skills attribute handout. After updating the cognitive skills attribute handout, the team reviewed the items from the operational forms again, with a focus on cognitive skill category 3. Then, the team reviewed the cognitive categories assigned to the 2021 operational form. Both reviews were used to further refine the updated MCAS cognitive skill attribute handout.

The mathematics team then used the updated cognitive skill attribute handout in conjunction with a careful analysis of the standards and the 2021 operational forms to develop the target distribution of score points by cognitive skill level. Finally, like ELA, the mathematics team developed training materials, including released assessment items as exemplars of each cognitive skill category assigned to items for each grade level. These examples are meant to be used during ADC orientation to explain how to assign cognitive skill categories based on cognitive demand and *not* item difficulty. Each set of exemplars is also

meant to provide a means to resolve any disagreements among ADC members. A consensus is reached by the ADC after discussions involving the descriptions and item exemplars.

Cognitive Skills Targets

The targets for score points by cognitive skill are given in the two tables below.

Table Z-1. Targeted Percentage of Score Points by Cognitive Skill Level in English Language Arts.

Grade	Cognitive Skill Level	Total Points	Percent of Score Points (+/-5%)	Score Points
3–4	I	44	5%	0–5
	II		70%	29–33
	III		25%	10–14
5	I	48	5%	0–5
	II		60%	26–31
	III		35%	14–17
6–8	I	50	5%	0–5
	II		60%	27–32
	III		35%	16–20
10	I	51	5%	0–5
	II		60%	28–33
	III		35%	16–21

Table Z-2. Targeted Percent of Score Points by Cognitive Skill Level in Mathematics.

Grade	Cognitive Skill Level	Total Points	Percent of Score Points	Score Points
3	I	48	25–40%	12–20
	II		55–65%	26–32
	III		6–15%	3–7
4–8	I	54	25–40%	13–22
	II		55–65%	29–35
	III		6–15%	3–8
10	I	60	25–35%	15–21
	II		55–65%	33–39
	III		7–20%	4 –12

Operational Implementation

Based on the cognitive skills attribute handout and exemplar items, all MCAS items are developed to a cognitive skill level, which is then recorded within the item metadata. The cognitive skill level for each item is first reviewed by DESE test developers along with standard alignment, content accuracy, readability, instructional worthiness, and other attributes of the item. The items are then reviewed by teacher committees, including the ADC, where they either accept or change the cognitive skill level.

Before ADC members review items, they are provided training, including definition of each cognitive skill level, along with examples of items that fall into each category. After items are field tested, the cognitive skill levels are again reviewed both by the test developers and by the teacher committees when they review the items with data after field testing. If the item is used as an operational item, the cognitive skill is again reviewed prior to the item being placed on the common form.

The targeted percentages of score points by cognitive skill are used by the contractor during the initial pull of each test for both ELA and math. The percentage of each cognitive skill is then reviewed and verified by DESE test developers during the test form construction period. If the targeted percentages are not reached, the contractor will focus on developing new items in the following year's development cycle to address any shortcomings.

References

Hess, K. (2013). *A Guide For Using Webb's Depth Of Knowledge With Common Core Standards.* Retrieved from The Ohio Department of Education: https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/How- to-Design-and-Select-Quality-Assessments/Webbs-DOK-Flip-Chart.pdf.aspx