

## Appendix D

### Best Practices in Mathematics Assessment

In reviewing the research on best practices in mathematics assessment, it becomes clear that practitioners should employ formats that are varied in nature, foster deep and rich thinking, and target the critical areas of skill mastery, concept development, and problem solving. Additionally, as outlined in the National Council of Teachers of Mathematics (NCTM) *Assessment Standards for School Mathematics*, quality assessment in mathematics should be an open and coherent process that enhances mathematical learning. It should simultaneously promote equity and the development of valid inferences about mathematical learning.<sup>23</sup>

As with assessment in any other discipline, mathematical assessment has a recursive relationship with instruction. When planning a unit, it is essential to reference the standards in order to determine what students need to learn and be able to do. Consequently, it is at this point that a teacher will be able to contemplate and design assessment models that they will administer both during and at the conclusion of the instructional roll-out. This pre-planning of assessment directs the path of instruction. The recursive nature of the relationship manifests itself through the use of formative and interim assessments. (e.g., screening, diagnostic, progress monitoring, and benchmark assessments). By utilizing the information garnered from these sources of data, a teacher is able to fine tune his/her instructional path to better serve the needs of all learners (e.g., creating small groups of students with similar needs for re-teaching, providing additional practice to others to cement concept development, or posing challenges to some that are ready for enrichment.) If the students' needs have been adequately addressed, their success

should be maximized when it comes time to evaluate student learning through the use of summative assessments (e.g., outcome measures).<sup>24</sup>

When designing mathematics assessments, it is necessary to consider the cognitive complexity of proposed tasks. In the Depth of Knowledge (DOK) Framework, Webb outlines four levels of cognitive demand – Recall, Skill/Concept, Strategic Thinking, and Extended Thinking. Level 1 cognition is characterized by simple recall. As illustrated in Table D.1, tasks of this nature require a rote response or the performance of a simple algorithm (e.g., *Find the next three terms . . .*). Level 2 DOK necessitates some type of decision making. The response to a prompt will not be automatic and will require some “mental processes” involving more than one step (e.g., *Draw the next figure in the following pattern . . .*). Moving up the spectrum to a Level 3 DOK, the quality of reasoning becomes more complex and demanding. Tasks of this variety require greater planning, abstraction, evidence and justification of thought. A student engaged in Level 3 cognition could be required to form a hypothesis or conjecture (e.g., *Find the next three terms in the pattern and determine the rule for. . .*). Lastly, the highest tier of DOK, Level 4, manifests itself in tasks that require an extended period of time utilizing complex thinking and planning. Level 4 tasks compel a student to make connections within a discipline and/or to other disciplines. More than likely, there are multiple solutions to a problem and multiple pathways for attaining a solution (e.g., *Find the next three terms in the pattern, determine the rule for finding the next number in the pattern, and make or find a model for . . .*).<sup>25</sup>

**Table D.1.**

<b>Recall &amp; Reproduction</b>  <b>(DOK 1)</b>	<b>Skills &amp; Concepts/ Basic Reasoning</b>  <b>(DOK 2)</b>	<b>Strategic Thinking/ Complex Reasoning</b>  <b>(DOK 3)</b>	<b>Extended Thinking &amp; Reasoning</b>  <b>(DOK 4)</b>
Find the next three terms in the following pattern:  $2/7, 4/7, 6/7\dots$	Draw the next figure in the following pattern:  	Find the next three terms in the pattern and determine the rule for the following pattern of numbers:  $1, 4, 8, 11, 15, 18, 22, 25, 29, \dots$	Find the next three terms in the pattern, determine the rule for finding the next number in the pattern, and make or find a model for the pattern:  $1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$ <sup>26</sup>

Bush and Leinwand offer some guidance in the selection and design of varied assessment items. When measuring skill attainment, items are best kept short necessitating a single correct answer. Generally involving simple recall or the use of a routine procedure, skill items require little or no context. Conceptual items can also be short in duration, but are not based on rote memorization and thus are non-routine. They are steeped in context requiring the production of some explanation or representation from the student. In relation to Norman Webb’s Depth of Knowledge levels, both of these types of items most likely fall within the scope of Level 1 or Level 2 cognition. Lastly, problem solving items capitalize on a student’s ability to apply their “mathematical power” in creating and using a plan. Such assessment items are context rich, require a sustained effort, are non-routine, and may have more than one correct solution.<sup>27</sup> The Depth of Knowledge for these items usually draws on Level 2

or Level 3 thinking, but in some complex situations reach to the top of the spectrum, Level 4. Students need to regularly encounter a balance of all of these types of assessment items in order to fully demonstrate the depth and breadth of their learning.<sup>28</sup>

When designing or selecting these varied assessment items, the infusion of writing in mathematics gives the teacher access to the inner thinking and reasoning of their students. Written evidence provides assistance in determining what the next instructional steps should be.<sup>29</sup> The use of writing is beneficial when assessing procedural or conceptual knowledge as well as reviewing a student’s justification for a solution path when problem solving.<sup>30</sup>

Finally, the role of the student in the assessment process cannot be overlooked. NCTM calls for a switch in focus from “students as the objects of assessment” to “students as active participants in

the assessment process.” As multiple measures are employed to assess a student’s “mathematical power,” a student is required to become more responsible for and reflective of their own learning.<sup>31</sup> When students engage in critical self-assessment of their own work and they gain frequent and targeted feedback from their teacher, they have a clearer understanding of expectations and their own mathematical learning. This combination fosters an environment that moves away from one that is judgment-oriented to one that focuses on the continued growth of all students.

### **Classroom Instructional Assessments for Mathematics**

A comprehensive assessment system provides multiple pieces of information on student achievement. Various types of assessments are required because they provide different types of information. The *Best Practices in Assessing Mathematics* section described what quality assessments in mathematics should look like. With respect to the purpose of an assessment, the outline that follows offers guidance as to what types of tools should be included within an LEA’s Comprehensive Assessment System.

**Screening Assessment(s)** ~ a type of interim assessment

- Used as a first alert or indication of being at risk for deficits in mathematics skills or concepts.
- Administered to ALL students *before* instruction.
- Quick and easy to administer to a large number of students.
- Correlated with content and/or instructional objectives germane to grade level performance.
- Rarely provide specific information needed to determine the most appropriate invention or target for instruction.

Mathematics screening instruments are broad in nature, so they are not able to address all facets of

grade level content or standards. With that said, their design should target the identification of common misunderstandings and should address the focal areas for that grade level. Items should assess factual and procedural knowledge as well as the application of concepts. Finally, they need to be reliable, have predictive validity, and be efficient in terms of administration and reporting.

Key questions that should be answered by the screening assessment(s):

- Which student(s) is experiencing mathematics difficulty?
- Which student is at risk for mathematics difficulty and in need of further diagnostic assessment(s) and/or additional interventions?

**Benchmark Assessments** ~ a type of interim assessment

- Used to chart growth in mathematics. Administered to all students.
- Determine if students are making adequate progress in overall performance towards standard(s).
- Typically administered at predetermined time (examples: end of a unit/theme, quarterly, etc.).

Key questions that should be answered by the benchmark assessments:

- What is the effectiveness of classroom instruction?
- Which student(s) need extra support to acquire a particular mathematics skill(s), concept(s), or standard(s)?
- How should groups be formed for classroom mathematics instruction?
- Which specific mathematics skills, concepts and/or standards need to be emphasized/ re-taught?

**Progress Monitoring** ~ a type of formative or interim assessment

- Used to determine next steps.
- Used during classroom mathematics instruction (may occur daily, weekly).
- Aligned to instructional objective.
- Can be used on an ongoing basis and may include teacher made-assessments, work samples, observational notes, and standardized or semi-structured measures of student performance.

Key questions that should be answered by the progress-monitoring assessments:

- How does the data articulate if the students “got it”?
- Does the lesson need to be re-taught to the whole class or just a few students?
- Who needs extra support?
- How is specific, constructive, and timely feedback provided to promote student learning or relearning of mathematics skills, concepts, or standards?

**Outcome Measures** ~ a type of summative assessment

- Used as a program or student evaluation.
- Used to indicate a student’s learning over a period of time and how proficient a student is towards meeting the grade level standards in mathematics.

Key questions that should be answered by the outcome assessments:

- To what degree has the student achieved the mathematics content standards?
- Is the assessment aligned to the state adopted mathematics standards?
- What information/data are provided and used to evaluate the effectiveness of the mathematics curriculum?
- Can decisions about selection, utilization of resources, materials and personnel be made with data collected from this mathematics assessment?

## **INTERVENTION ASSESSMENTS**

**Diagnostic Assessment(s)** ~ a type of interim assessment

- Used to gain an in-depth view of a student’s mathematics profile.
- Administered to students who have been identified as at-risk of not achieving grade level mathematical proficiency during the screening process.
- Often are individually administered so observations of behaviors may also be included.

Diagnostic assessments are used to determine gaps in student learning. They provide evidence to make inferences with respect to instructional interventions.

Key questions that should be answered by the diagnostic assessments:

- What are a student’s strengths in mathematics?
- What are a student’s weaknesses in mathematics?
- What are the specific areas of need for the student?
- Are other students exhibiting similar mathematics profiles?
- How should mathematics intervention groups be formed?

**Progress Monitoring of Intervention** ~ a type of formative or interim assessment

- Used to chart growth towards benchmark/goal/standard.
- Used for students who have intervention services in mathematics.

Key questions that should be answered by the progress-monitoring assessments when using an intervention:

- Is the individual student progressing at a sufficient rate to achieve the goal?
- Are instructional revisions needed in order for sufficient progress to be made towards the student’s benchmark/goal/standard?
- Has this intervention been proven effective in improving students’ mathematics skills?

## Sources

---

- <sup>1</sup> Trimble, S. (2003). NMSA research summary #20: What works to improve student achievement. Westerville, Ohio: National Middle School Association. Retrieved from: [http://www.nmsa.org/portals/0/pdf/publications/On\\_Target/achievement/achievement\\_4.pdf](http://www.nmsa.org/portals/0/pdf/publications/On_Target/achievement/achievement_4.pdf)
- <sup>2</sup> Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- <sup>3</sup> Popham, W.J. (2008) Transformative Assessment. Alexandria, VA: ASCD.
- <sup>4</sup> Wiliam, D. (2009). *Assessment for Learning: Why, what, and how?* London: Institute of Education University of London.
- <sup>5</sup> Rowe, M. B. (1974). Wait time and rewards as instructional variables, their influence on language, logic and fate control: Part one-wait-time. *Journal of Research in Science Teaching*, 11, 81-94. doi: 10.1002/tea.3660110202
- <sup>6</sup> Shavelson, R. J. (2006). *On the integration of formative assessment in teaching and learning with implications for teacher education*. Paper prepared for the Stanford Education Assessment Laboratory and the University of Hawaii Curriculum Research and Development Group. Available at [www.stanford.edu/dept/SUSE/SEAL](http://www.stanford.edu/dept/SUSE/SEAL)
- <sup>7</sup> Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). The role of interim assessments in a comprehensive assessment system: A policy brief. Retrieved from <http://www.achieve.org/files/TheRoleofInterimAssessments.pdf>
- <sup>8</sup> Perie, M., Marion, S., & Gong, B. (2007). A framework for considering interim assessments. Washington, D.C.: National Center for the Improvement of Educational Assessment. Retrieved from [http://www.nciea.org/publications/ConsideringInterimAssess\\_MAP07.pdf](http://www.nciea.org/publications/ConsideringInterimAssess_MAP07.pdf).
- <sup>9</sup> The Council of Chief State School Officers (2008). Interim assessment practices and avenues for state involvement TILSA SCASS interim assessment subcommittee. Retrieved from [http://www.ccsso.org/Documents/2008/Interim\\_Assessment\\_Practices\\_and\\_2008.pdf](http://www.ccsso.org/Documents/2008/Interim_Assessment_Practices_and_2008.pdf)
- <sup>10</sup> American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: Author.
- <sup>11</sup> Trochim, W. (2006). Research methods knowledge base: Reliability & validity. Retrieved from <http://www.socialresearchmethods.net/kb/relandval.php>
- <sup>12</sup> National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D.C.: National Academy Press.
- <sup>13</sup> Edwards, P.A., Turner, J.D., & Mokhtari, K. (2008). Balancing the assessment of learning and for learning in support of student literacy achievement. *The Reading Teacher*, 61(8), 682-684. doi: 10.1598/RT.61.8.12
- <sup>14</sup> Rhodes, L. K., & Shankin, N. L. (1993). *Windows into Literacy: Assessing learners K-8*. Heinemann: Portsmouth, NH.

- <sup>15</sup> Torgesen, J.K. (2006). A comprehensive K-3 reading assessment plan: Guidance for school leaders. Portsmouth, NH. RMC Research Corporation, Center on Instruction.
- <sup>16</sup> Webb, N.K. (2002) Depth-of-Knowledge Levels for Four Content Areas. Wisconsin Center for Educational Research.
- <sup>17</sup> Olinghouse, N. G. (2009). Writing assessment for struggling learners. *Perspectives in Language and Literacy*, Summer, 15-18.
- <sup>18</sup> NCTE Writing Assessment: A Position Statement (2009).
- <sup>19</sup> NCTE Writing Assessment: A Position Statement (2009).
- <sup>20</sup> Routman, R. (2005). *Writing Essentials: Raising Expectations and Results While Simplifying Teaching*. Portsmouth, NH: Heinemann.
- <sup>21</sup> Strickland, K. & Strickland, R. (2000). Making Assessment Elementary. Portsmouth, NH: Heinemann.
- <sup>22</sup> Newkirk, T. & Kent, R. (2007). *Teaching the Neglected "R": Rethinking Writing Instruction in the Secondary Classrooms*. Portsmouth, NH: Heinemann.
- <sup>23</sup> National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.
- <sup>24</sup> American Federation of Teachers. (2008). Making classroom assessments work for you: Training Materials. Washington, DC: Author.
- <sup>25</sup> Webb, N.& Collins, M. (2008). *Depth of knowledge: Descriptors, examples and question stems for increasing depth of knowledge in the classroom*. Kirksville, MO: Truman State University.
- <sup>26</sup> Kentucky Department of Education (2007). Support materials for core content for assessment version 4.1: Mathematics. Retrieved from <http://www.education.ky.gov/kde/instructional+resources/curriculum+documents+and+resources/core+content+for+assessment/core+content+for+assessment+4.1/content+specific+core+content+for+assessment+dok+support+materials.htm>
- <sup>27</sup> Bush, W.S. & Leinwand, S.(Eds.). (2000). *Mathematics assessment: A practical handbook for grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.
- <sup>28</sup> Sutton, J. & Krueger, A. (Eds.). (2002). *ED thoughts: What we know about mathematics teaching and learning*. Aurora, CO: Mid-continent Research for Education and Learning.
- <sup>29</sup> Ibid.
- <sup>30</sup> Bush, W.S. & Leinwand, S.(Eds.). (2000). *Mathematics assessment: A practical handbook for grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.
- <sup>31</sup> Lesh, R. & Lamon, S.J. (1992). *Assessment of authentic performance in school mathematics*. Washington, DC: American Association for the Advancement of Science.
- <sup>32</sup> National Science Teachers Association. (2010). Science education assessments - NSTA's official position on assessment. Arlington, VA: Author. Retrieved from <http://www.nsta.org/about/positions/assessment.aspx>

- <sup>33</sup> Keeley, P. (2008). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning*. Thousand Oaks, CA: Corwin Press.
- <sup>34</sup> Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Education: Principles, Policy and Practice*, 5, (1), 7-71.
- <sup>35</sup> Porter, A. (1993). *Brief to policymakers: Opportunity to learn*. Center on Organization and Restructuring of Schools. Retrieved from [http://www.wcer.wisc.edu/archive/cors/brief\\_to\\_principals/BRIEF\\_NO\\_7\\_FALL\\_1993.pdf](http://www.wcer.wisc.edu/archive/cors/brief_to_principals/BRIEF_NO_7_FALL_1993.pdf)
- <sup>36</sup> National Council for the Social Studies. (2008). A vision of powerful teaching and learning in the social studies: Building effective citizens. *Social Education*, 72 (5), 277-280.
- <sup>37</sup> Ibid.
- <sup>38</sup> Alleman, J. & Brophy, J. (1999). The changing nature and purpose of assessment in the social studies classroom. *Social Education*, 65 (6), 334-337.
- <sup>39</sup> Ibid.