The Challenges of Measuring Misconceptions in Middle Grades Statistics
Jessica Masters and Lisa Famularo
Measured Progress Innovation Lab

ABSTRACT

The Diagnostic Statistics Assessment (DSA) project is developing a formative, cognitively diagnostic assessment system designed to measure three misconceptions in middle grades statistics. The DSA consists of diagnostic score reports, and instructional activities designed to help students reorganize their understandings. The DSA researchers have conducted item development and item revision based on expert review and cognitive labs. Based on these activities, the authors describe the challenges and lessons learned in developing items to target statistical misconceptions at this grade level.

PROJECT OVERVIEW

Purpose: Develop a formative, cognitively diagnostic assessment system for middle grades teachers to identify and address these three misconceptions their students hold in statistics.

Misconceptions: The DSA targets three misconceptions grounded in cognitive science research. The authors use the term misconception broadly to reflect the idea that misconceptions often represent knowledge that is productive in some contexts but is underdeveloped or generalized:

• Datasets as Entities: Students do not perceive a dataset and its own characteristics, unique from those of its individual data points.
• Comparing Datasets: Students do not compare datasets based on representative or summary measures, but rather based on single data points, cutoff points, or slices.
• Overreliance on the Mean Procedure: Students reorganize their understandings. The DSA researchers have conducted item development and item revision (expert review and cognitive labs), the DSA researchers have conducted item development and item revision (expert review and cognitive labs), the DSA researchers have conducted item development and item revision (expert review and cognitive labs), the DSA researchers have conducted item development and item revision (expert review and cognitive labs), the DSA researchers have conducted item development and item revision (expert review and cognitive labs).

Research and Development: The DSA items will undergo multiple iterative rounds of research and revision:

• Expert review to collect evidence of validity based on test content
• Cognitive labs using a “think aloud” protocol to collect evidence of validity based on response process
• Field test to collect evidence of validity based on internal structure and the relationship to other variables

SESSION PURPOSE

Throughout the first two phases of item development and revision (expert review and cognitive labs), the researchers encountered challenges unique to the measurement of statistical knowledge, particularly in the middle grades. Four main challenges are presented, along with recommendations for action. The goal is to inform the development of future assessments and instructional materials related to statistics in the middle grades.

CHALLENGE: AMOUNT OF CONTEXT

How much information should be provided in order to create a real-world context without obscuring the statistical construct being measured?

The phrase “solve real-world and mathematical problems” is used repeatedly throughout the Common Core. The use of context in mathematics assessment is intended to both improve students’ ability to demonstrate their understanding of mathematics and to measure a deeper understanding of mathematical principles. Context is particularly important in statistics. Decisions about data collection, analysis of data, and the interpretation of results are heavily dependent on the research questions which are grounded in a particular context or situation.

Choosing an appropriate level of context for an item can be particularly challenging when the goal is to measure the more mathematical aspects of statistical knowledge. Asking the student to find the mean of a list of numbers encourages students to apply the mean procedure in a vacuum. Students should be encouraged to consider the context of the situation and the data to evaluate whether the mean is an appropriate measure. But the inclusion of additional contextual information might place an unnecessary burden on students that could be distracting and/or lead to an inaccurate assessment of their abilities.

How much data and information to provide is also important and needs to be made with careful consideration of the grade level of the student. It is important to find a balance in providing enough data so that students can find the answer they need, but too much data can overwhelm or confuse them. All possibilities and “what if” questions that might be asked by a statistician might not be considered by an average middle school student. They might, however, be considered by an above average student.

CHALLENGE: MEANINGFUL, AGE-APPROPRIATE, AND REALISTIC CONTEXTS

The context must be meaningful for all students who will complete the assessment. Certain topics might be familiar to most students, but if they cause issues for one group, they can be avoided. For example, the use of context in mathematics assessment is intended to both improve students’ ability to demonstrate their understanding of mathematics and to measure a deeper understanding of mathematical principles. Context is particularly important in statistics. Decisions about data collection, analysis of data, and the interpretation of results are heavily dependent on the research questions which are grounded in a particular context or situation.

Choosing an appropriate level of context for an item can be particularly challenging when the goal is to measure the more mathematical aspects of statistical knowledge. Asking the student to find the mean of a list of numbers encourages students to apply the mean procedure in a vacuum. Students should be encouraged to consider the context of the situation and the data to evaluate whether the mean is an appropriate measure. But the inclusion of additional contextual information might place an unnecessary burden on students that could be distracting and/or lead to an inaccurate assessment of their abilities.

How much data and information to provide is also important and needs to be made with careful consideration of the grade level of the student. It is important to find a balance in providing enough data so that students can find the answer they need, but too much data can overwhelm or confuse them. All possibilities and “what if” questions that might be asked by a statistician might not be considered by an average middle school student. They might, however, be considered by an above average student.

CHALLENGE: USE OF GRAPHS

How can statistical knowledge be measured without being confounded by the use of graphs and histograms?

The assessment of statistical knowledge often requires students to analyze data. Most meaningful datasets will be large enough to require a graphical presentation, as opposed to a list of raw values. The student must be able to interpret the graph in order to access the data required to respond to the statistical inquiry. If a student is unable to interpret the graph (e.g., does not understand what the axes represent, incorrectly reads the scale, etc.), the student will not be able to successfully address the statistical question. Thus, the ability to interpret a graph or histogram can serve as an indicator of the demonstration of statistical knowledge. Histograms are historically the most challenging type of graph to interpret but avoiding them places further restrictions on measuring meaningful contexts.

CHALLENGE: STATISTICAL LANGUAGE

How can statistical knowledge be measured without being confounded by statistical language?

Because the Common Core has placed a greater emphasis on statistics and probability than previous standards and most of the statistical content is introduced in the Grades 6-8 standards, many middle grades students will be introduced to statistics for the first time. And many of their teachers will be teaching it for the first time. DSA cognitive labs revealed that putting the student in the role of an adult decision maker increased the cognitive load thus possibly preventing the student from demonstrating his/her knowledge. In other cases, there are no meaningful contexts for which middle school students truly get to act as a decision maker!

A final challenge in finding meaningful and realistic contexts is avoiding the use of decimals in datasets. Cognitive labs revealed that students’ difficulties with decimals can be a confounding factor when attempting statistical calculations. But finding meaningful datasets that use only whole numbers is a challenge.

Another challenge stems from use of the term average. Is average synonymous with mean? Or could it also mean weighted average? This issue is further complicated by the fact that average both in a mathematical and everyday term. Many experts recommend not using the term average and rather using the statistical term for what measure is required. However, what if the assessment seeks to determine whether students can choose which measure should be used?

RECOMMENDATIONS

• When assessing statistical knowledge, provide students with context, even when targeting the more mathematical aspects of statistical knowledge. Context should include the statistical question being asked, not just a description of what the numbers represent. This is a better reflection of how students are encouraged to apply statistical rigor.
• In the classroom, instruction should focus on meaningful applications of mathematical processes within statistics. Assessments should mirror that instructional approach.
• Item writers must strike a balance between providing enough contextual information to create a realistic but not overly complicated problem situation. Scenarios should support, not distract from, the statistical concepts.
• Teachers must understand more than the topic being assessed, and abilities being assessed. They must have a strong understanding of what scenarios and language are developmentally appropriate for students at the targeted grade level.
• Item writers need feedback from various sources to find the appropriate balance of item context. Sources should include both statistical experts and teachers at varying levels of proficiency.
• Teachers should begin to use appropriate statistical terms in early grades and use these terms consistently throughout elementary and middle school. Assessments can then reflect the language used in the classroom. In the interim, caution should be employed on the use of seemingly obvious statistical language in assessments.
• Careful attention should be paid to the amount of information, layout of information, and language used in labeling graphs. Keys and scales should be included to help students understand and interpret graphs. When possible, similar layout and formatting should be used for all items within a single assessment.