PEARSON EDUCATIONAL MEASUREMENT Bulletin

PEARSONEDMEASUREMENT.COM

MARCH 2007 • ISSUE 3

Griddable Items: Beyond Multiple-Choice

Multiple-choice test items along with "fill-in-the-bubble" answer sheets are mainstay formats in testing—formats that are prevalent for practical reasons. Pearson Educational Measurement (PEM), who pioneered scanning technology in 1953, has led the technological innovations that have helped refine these formats.

Multiple-choice test items are practical in their efficiency to be administered and scored in large volumes over a relatively short period of time. In addition, the construction of multiple-choice test items is well understood. Over the years, measurement experts have offered specific guidelines for creating high-quality questions and response options.

But today in the testing industry, there is a heightened need to find alternatives to the multiple-choice format: alternatives that can offer more diagnostic information while still offering practical efficiencies and strong measurement qualities. PEM is performing research on one alternative format—the griddable test item. Griddable items are a special case of constructed response items, which require the response to be generated by the examinee instead of selecting a response from a set of options. Under the griddable item format, an examinee grids in their response, usually a number, using several columns of response bubbles. By using response bubbles, griddable items offer the same scanning efficiencies as multiple-choice.

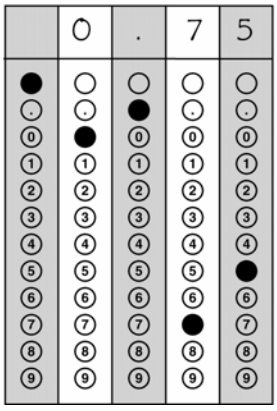
An example of a griddable test item and the completed response bubbles are shown next.

1 Debi and some students from her geometry class went outside to measure their shadows. Debi measured Sean's and Karyn's shadows and recorded the information in the table below.

SHADOW MEASUREMENTS

Name	Height (x)	Shadow Length (y)
Sean	159 cm	106 cm
Karyn	167 cm	112 cm

Debi plotted this data on a coordinate grid. What is the slope of the line containing these two points?



Before considering this alternative item format, there must be a good understanding of student performance on griddable items. Current research is sparse but can help answer some important questions with regard to griddable items.

PEARSON EDUCATIONAL MEASUREMENT Bulletin
PEARSONEDMEASUREMENT.COM
MARCH 2007 • ISSUE 3

What range of constructs can be measured by griddable items? Griddable items may be able to elicit a broader range of cognition than multiple-choice items[1], but the conclusions of research are mixed. One study[2] concluded that the griddable item format requires students to analyze and solve problems without being influenced by multiple-choice response options. Another study[3] concluded that the same reasoning skills are used by students on both multiple-choice and griddable items. These mixed results may be attributed to the way multiple-choice items were constructed. Griddable items may well measure a broader range of cognition than multiple-choice items if the multiple choice response options were not an accurate reflection of the errors actually made by the students.

Do griddable items tend to be more challenging for students than conventional multiple-choice items? Research studies have found that griddable items were more difficult than multiple-choice items[4,5,6]. One study[4] directly compared the same items in both griddable and multiple-choice formats and found that some items that were relatively easy in multiple-choice format were relatively difficult in griddable format. In addition, the more challenging nature of griddable items may influence students to skip items. A study[5] found that grid-in items were omitted with a higher frequency than multiple-choice items.

Do students from different focal groups (e.g., female or African American students) perform differently on griddable items compared to a reference group (e.g., male or Caucasian students)? Research on this issue is fairly clear. For well constructed items, focal groups perform similarly to the reference group on griddable items[7,8]. However, as with multiple-choice items, all items constructed using the griddable item format should be submitted to review for sensitive content.

Why might the griddable item format be adopted for at least some of the items on a test? The griddable item format may offer some advantages over the multiple-choice item format. First, the use of griddable items on tests may encourage deeper, more authentic student learning and discourage student memorization[1]. For example, one study[2] reported that teachers and students believe that griddable items reflect mathematical problem solving better than multiple-choice items. Second, performance on griddable items may offer educators more diagnostic information. One study[4] claims that student performance on griddable items offers better insight into the specific skills students possess than does student performance on multiple-choice items. However, PEM is doing more research to substantiate claims of advantages for the griddable item format over the multiple-choice item format.

PEARSON EDUCATIONAL MEASUREMENT Bulletin

PEARSONEDMEASUREMENT.COM

MARCH 2007 • ISSUE 3

PEM's research is investigating the griddable item format for its measurement quality as well as its diagnostic advantages. PEM is presenting research on the griddable item format at the CCSSO National Conference on Large-Scale Assessment, June 17-20, 2007. The 120-minute symposium is titled, "Gridded-Response Items: Should They Be Used in High-Stakes Testing?" Please check the conference program for time and location.

— Paul Nichols, Ph.D.

[1] Martinez, Michael E. (1999). Cognition and the Question of Test Item Format. Educational Psychologist, 34, 207-218.

[2] Braswell, J., & Jackson, C. (1995). An introduction of a new free-response item type in mathematics. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA, April 19-21.

[3] Katz, I. R., Friedman, D. E., Bennett, R. E., & Berger, A. E. (1996). Differences in strategies used to solve stem-equivalent constructed-response and multiple-choice SAT-mathematics items. College Board Report (95-3).

[4] Bridgeman, B. (1992). A comparison of quantitative questions in open-ended and multiple-choice formats. Journal of Educational Measurement, 29, 253-271.

[5] Hombo, C. M., Pashley, K., & Jenkins, F. (April, 2001). Are grid-in format items usable in secondary classrooms? Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, April 11-13, 2001.

[6] Schmitt, A.P., & Crone, C.R. (1991). Alternative Mathematical Aptitude Item Types: DIF Issues. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, April 4-6 1991.

[7] Burton, N. (1996, Winter). Have changes in the SAT affected women's mathematical performance? Educational Measurement: Issues and Practice, 5-9.

[8] Garner, M., & Englehard, G. (1999). Gender differences in performance on multiple-choice and constructed response mathematical items. Applied Measurement in Education, 12, 29-51.

Pearson Educational Measurement (PEM) has been a trusted partner in district, state and national assessments for more than 50 years. As a pioneer and the largest comprehensive provider of educational assessment products, services and solutions, PEM helps states and large school districts meet the requirements of education reform while using assessment and research to promote learning.

More information is available at www.PearsonEdMeasurement.com.