EDUCATIONAL DESIGNER
JOURNAL OF THE INTERNATIONAL SOCIETY FOR DESIGN AND DEVELOPMENT IN EDUCATION

Periodic Assessments and Diagnostic Reports

Case Studies in Mathematics and Literacy Intervention Programs

Betsy Taleporos
Former Director of Assessment, America's Choice

Abstract

This paper discusses the formative use of periodic assessments as they were developed and are in use by America's Choice Pearson in its mathematics and language arts intervention programs. It is a practical case study of the use of design principles in creating assessments that are useful for classroom teachers and, by the nature of their design, provide diagnostic information that is instructionally relevant. The use of these measures varies with the program but all of them are designed to highlight misconceptions or common error patterns. It is important to recognize that misconceptions occur in both content domains, as they do in other domains. Uncovering misconceptions or error patterns offers tremendous insight into a formative use of assessments, since the reasons behind answering a question incorrectly can directly inform instructional practice. This approach is also underscored by some of the suggestions in the lead article in this issue of ED.

Overview

Today's assessment landscape is changing, but remains dominated by large-scale testing which, as indicated by the lead article in this issue, is fraught with problems that are not always in sync with the needs of the classroom teacher. The current state test reports give information that is generally broader in scope than the information a classroom teacher needs to help students improve in the learning expected by the instruction given to them directly and specifically.

The nature of state test reports do not lend themselves to diagnosis or focusing on specific needs of students in a way that lets teachers plan to meet those needs in their day-to-day practice. The information is not provided in a timely manner, often received months after students take tests. Where teachers can look at the results of their current, not last year's, class, the information is generally too broad

to be of practical use. Further, the type of tasks provided for students to work on in most state testing situations rarely tap deep understanding .

While much is wrong with the current system, the new consortia for assessing the Common Core State Standards are making attempts to correct some of the current flaws, including enhanced item types and an emphasis on formative assessment during the school year. Currently, for both consortia, the formative assessments are optional, and outside the formal accountability measurement, but their value is clearly recognized. Whether the fact that they are optional, and don't count in a final accountability score, will weaken their impact is yet to be seen.

The new item types, however, are bound to make an impact on classroom instruction, where so much time is spent on prepping for the annual accountability tests. If those tests are significantly different than the ones currently used by most states, then the impact will undoubtedly be positive. Nonetheless, the system is still plagued with the issues surrounding the need for continual feeds of information on how well students are learning what they are being taught. The need for formative assessment will still be as critical as it is now with the current individual state testing systems.

Classroom assessments have their own set of problems as well. Teachers receive little guidance in test construction in their pre-service training or their continual professional development. The resulting assessments may not be as rigorous as needed, and the quality of the items included may not be optimal. Nonetheless, they are a reflection of what is valued by the teacher, a measure of the intended curriculum as well as of the enacted curriculum.

This paper contains figures which llustrate some of the features of both the mathematics and language arts assessments. The figures also include screenshots of parts of the on-line reports, which are at the heart of the assessment. Because real time access to the reports is proprietary, only screenshots could be shown in this paper.

Mathematics Navigator

The Program

Mathematic Navigator is an intervention program, designed for students who need some additional time and focused teaching in specific areas of mathematics. There are 26 modules in this program, each focusing on a different targeted area of mathematics, such as Place Value, Fractions, Data and Probability, Exponents, Expressions and Equations, Rational Numbers, to name but a few.

The Assessments

Each of these modules has a pretest and a posttest, as well as checkpoint assessments. There is also an omnibus screener for each grade level to help

determine students' needs for particular modules. Figure 1 shows the assessments that are part of the Mathematics Navigator program. Figure 2 lists the reports and shows the levels of aggregation possible for each of them. It also shows the purpose of each of the reports.

Figure 1: Mathematics Navigator assessments

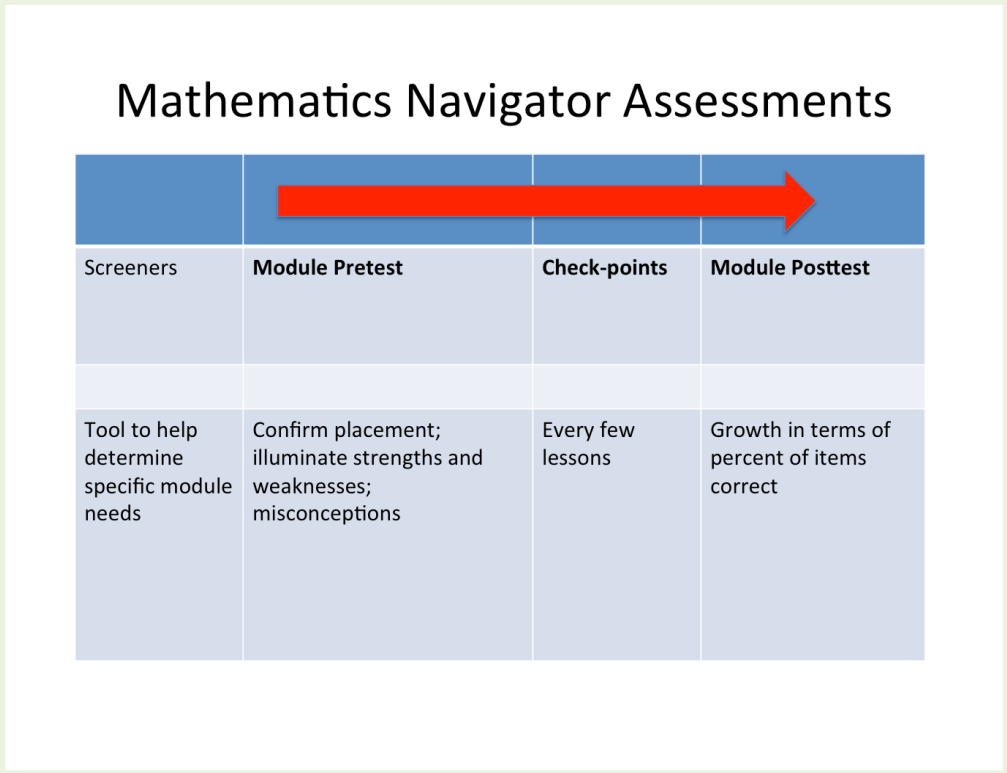
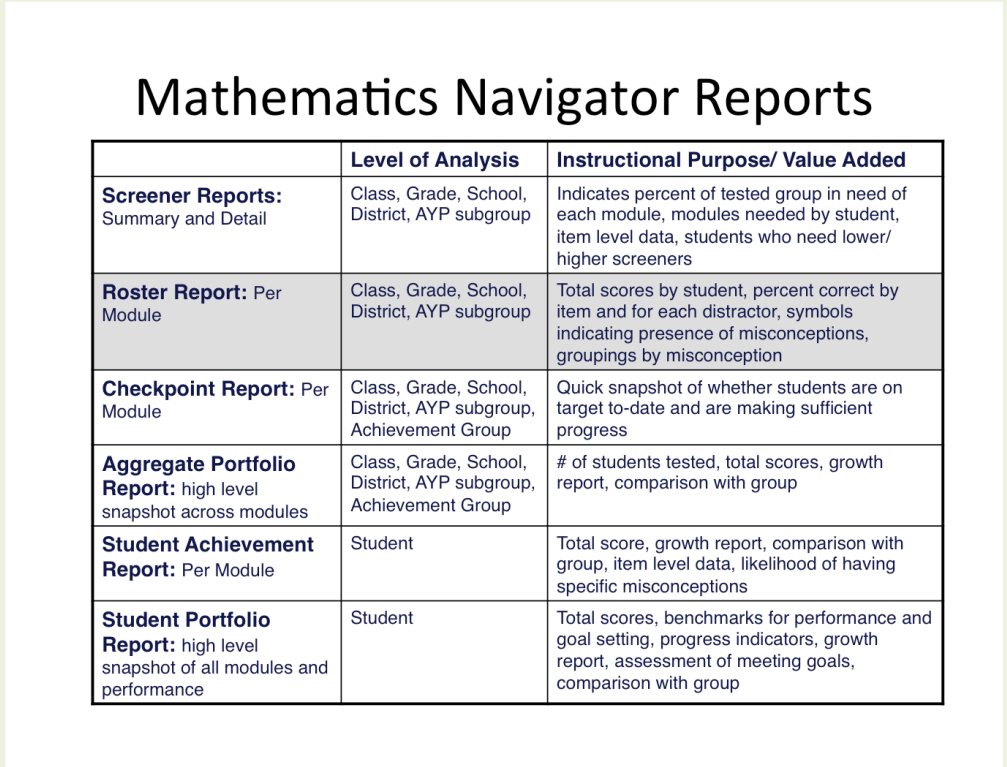


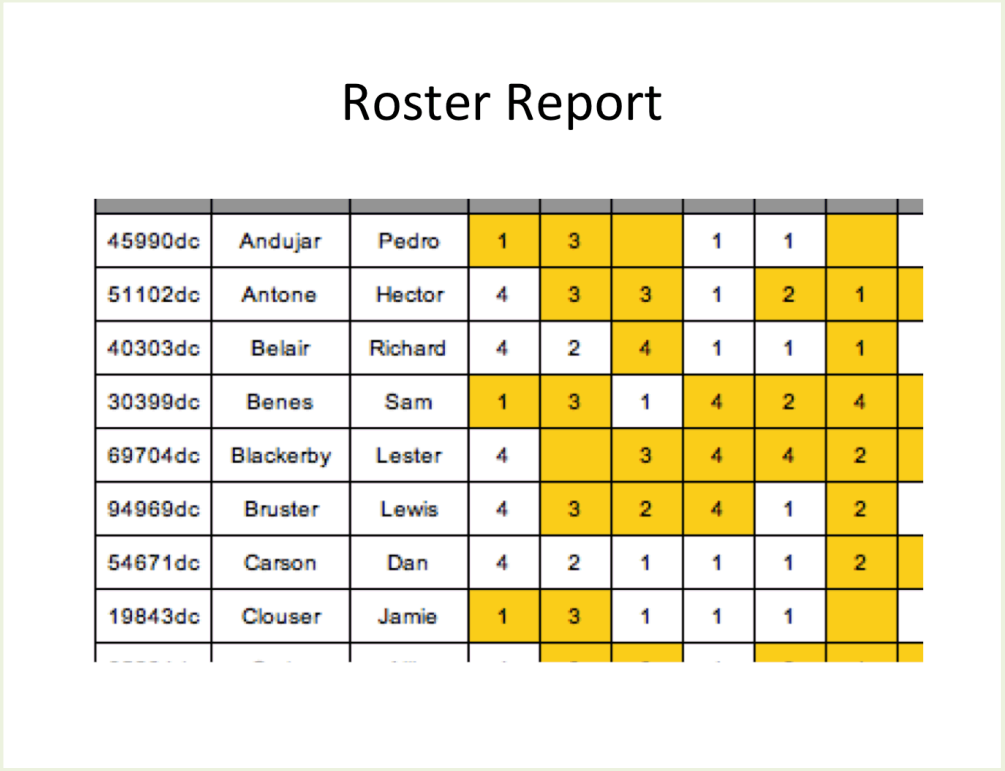
Figure 2: Mathematics Navigator reports



	Level of Analysis	Instructional Purpose/ Value Added
Screener Reports: Summary and Detail	Class, Grade, School, District, AYP subgroup	Indicates percent of tested group in need of each module, modules needed by student, item level data, students who need lower/ higher screeners
Roster Report: Per Module	Class, Grade, School, District, AYP subgroup	Total scores by student, percent correct by item and for each distractor, symbols indicating presence of misconceptions, groupings by misconception
Checkpoint Report: Per Module	Class, Grade, School, District, AYP subgroup, Achievement Group	Quick snapshot of whether students are on target to-date and are making sufficient progress
Aggregate Portfolio Report: high level snapshot across modules	Class, Grade, School, District, AYP subgroup, Achievement Group	# of students tested, total scores, growth report, comparison with group
Student Achievement Report: Per Module	Student	Total score, growth report, comparison with group, item level data, likelihood of having specific misconceptions
Student Portfolio Report: high level snapshot of all modules and performance	Student	Total scores, benchmarks for performance and goal setting, progress indicators, growth report, assessment of meeting goals, comparison with group

The testing reports are essential online tools for the teacher to use in implementing the program. The reports focus on diagnosis and performance levels.

Figure 3: Roster report



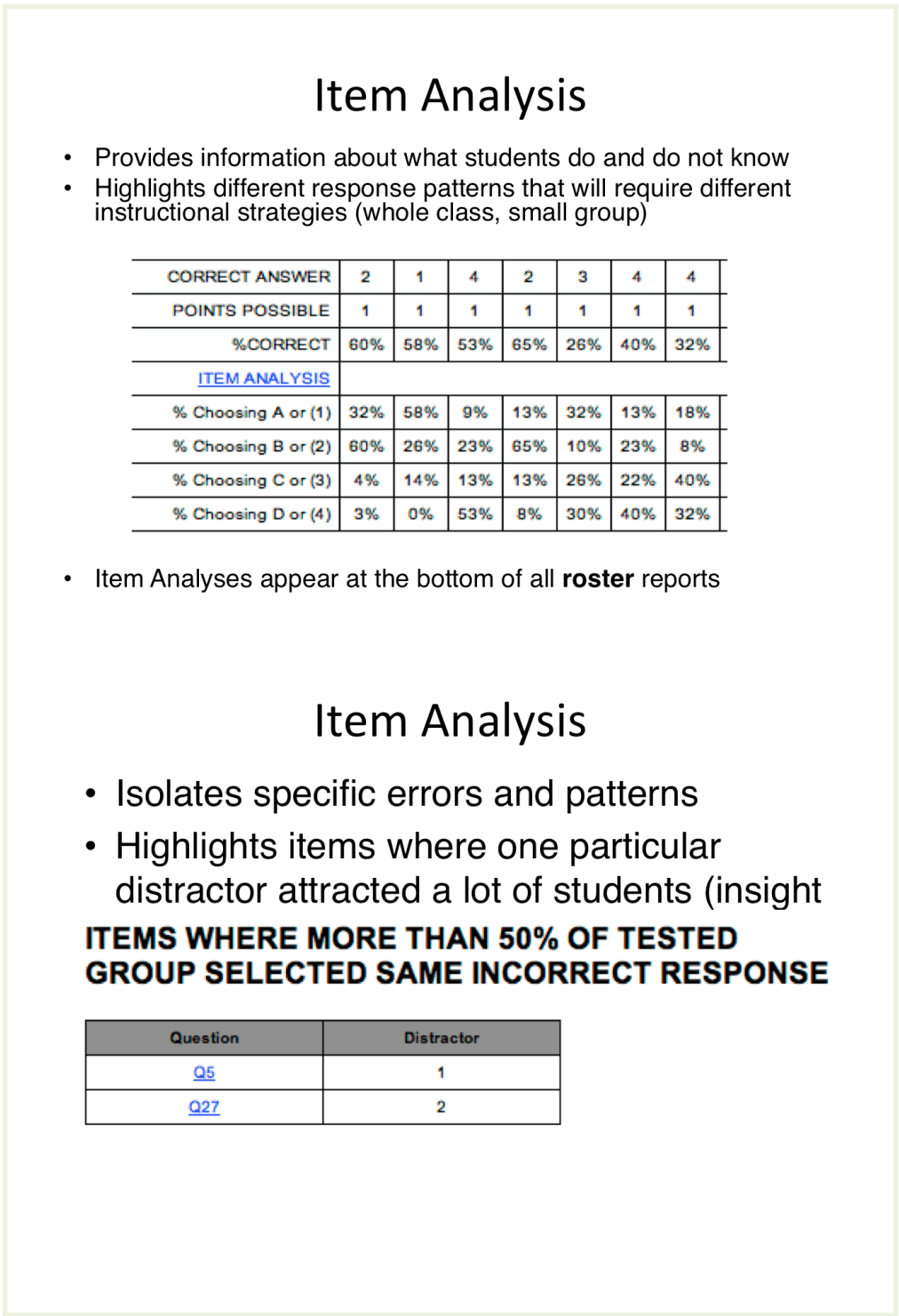
Diagnostic Reports

One report, called a roster report, shows the answer that each student gave to each question, and also shows a listing of the misconceptions that these answers show the student to have. Each question number on the report is hyperlinked so that the teacher can click on it and see the actual question and the answer choices to see what specific choice a student has made. Part of the roster report is shown in Figure 3. It shows how each student's choice is provided, and is shaded yellow if incorrect. The teacher can get a bird's eye view of how well a whole class did on an assessment simply by looking at the proportion of item choices that are shaded yellow, but also by looking at quantitative information on the roster report itself.

Item Analysis

The roster report also shows the percent of student getting the item correct, and the percent choosing each answer choice. The report also highlights individual questions where the majority of students got the same wrong answer. These bits of information are useful to teachers to get a broad view of the needs of the whole group of students in the mathematics navigator class. The item analysis information is shown in Figure 4.

Figure 4: Item evaluation



Item Analysis

- Provides information about what students do and do not know
- Highlights different response patterns that will require different instructional strategies (whole class, small group)

CORRECT ANSWER	2	1	4	2	3	4	4
POINTS POSSIBLE	1	1	1	1	1	1	1
%CORRECT	60%	58%	53%	65%	26%	40%	32%
ITEM ANALYSIS							
% Choosing A or (1)	32%	58%	9%	13%	32%	13%	18%
% Choosing B or (2)	60%	26%	23%	65%	10%	23%	8%
% Choosing C or (3)	4%	14%	13%	13%	26%	22%	40%
% Choosing D or (4)	3%	0%	53%	8%	30%	40%	32%

- Item Analyses appear at the bottom of all **roster** reports

Item Analysis

- Isolates specific errors and patterns
- Highlights items where one particular distractor attracted a lot of students (insight

ITEMS WHERE MORE THAN 50% OF TESTED GROUP SELECTED SAME INCORRECT RESPONSE

Question	Distractor
Q5	1
Q27	2

Test Design – Focus on Misconceptions

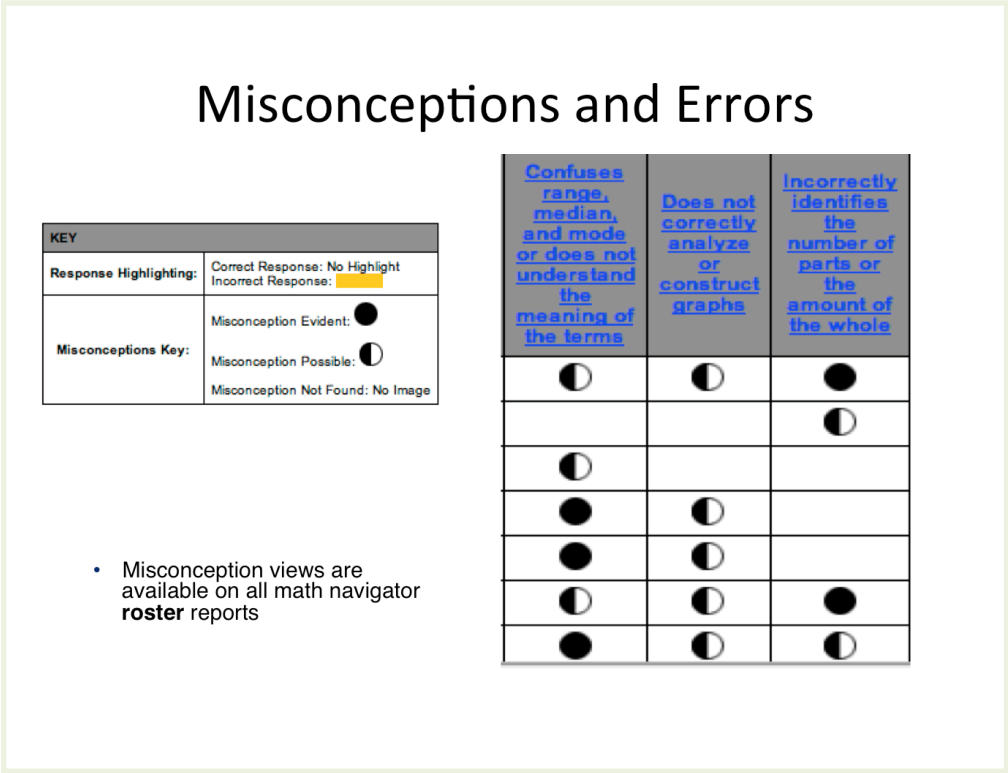
The tests are designed in a very purposeful way. The items are all multiple choice, measuring key concepts taught in the module. The wrong answer choices are coded to common misconceptions so a student's pattern of answer choices can be used to describe the misconceptions that they have. For each misconception, there are at least four opportunities for a student to choose an option that reflects it. If the student typically chooses the wrong answers that reflect the misconception, the

report will show that they have that particular misconception.

As a design issue, the minimal number of four opportunities was chosen in somewhat of an arbitrary fashion, based on experience and industry standard approaches. This number is thought to provide stable enough estimation, given the chance to see a recurring pattern of selecting errors that reflect the given misconception. If a student makes a selection of a given misconception at least 75% of the time, we can be fairly confident that they have the given misconception. If it is chosen between 50 and 74% of the time, we conclude that they may have the misconception, but we are not as sure as we are when they systematically select the wrong answer with that misconception. Anything less than 50% does not permit us to make a conclusion about the systematic reflection of a given misconception.

Using the percent of times the student picks the answer reflecting a given misconception, the report will show either that the student definitely has the misconception, possibly has the misconception, or that there is no evidence of a pattern indicating that the student has the misconception. Figure 5 shows this report.

Figure 5: Misconceptions and errors



Grouping of Students

The reports also provide a listing of students by their misconception patterns that are often useful to teachers in setting up small group instruction. This information is used by teachers to have a diagnostic understanding of their students, and can be used to guide instruction for them. Teachers can group students together who

have similar misconceptions, or can group a student with a given misconception with another student who understands the misconception.

Checkpoint Assessments

The checkpoint assessments are provided several times over the course of the module. Each includes a debugging activity in which the students are asked to review each wrong answer and determine the thought pattern that would have led to the choice of that wrong answer. This is an additional design feature that enhances the diagnostic value of the checkpoint assessments as the discussion focuses on the thought patterns that exemplify misconceptions.

Figure 6 shows a report of the checkpoint assessments. The number of correct answers is transformed by a predetermined cutpoint, to indicate that the student is doing well (shaded green), may be having some difficulties (shaded yellow) or is have a great deal of trouble (shaded red). The cutpoints vary with the checkpoint assessments, determined by expert judgment for each one. While not scaled together in a psychometric analysis, the use of the judgment methodology simply indicates the student's status on the given checkpoint, and whether their relative status has changed from one checkpoint time to another.

Figure 6: Checkpoint report

Math Navigator Checkpoint Report

Class report for	DC, Jane Doe School District / Jane Doe School District / Kennedy MS / Math Navigator
Test Group	Understanding Fractions
Grade Levels	All
AYP Categories	All

Key	
Color Codes:	: Indicates student is on target and making sufficient progress. : Indicates student may not be on target. : Indicates student is not on target and is not making sufficient progress.

Understanding Fractions Math Navigator Checkpoint Report		
Student	Checkpoint 1	Checkpoint 2
Clouser, Jamie		
Corby, Allie		
Deak, Neil		
Esqueda, Javier		
Harger, Harriett		
Labriola, Kelly		
Mazzariello, Ted		

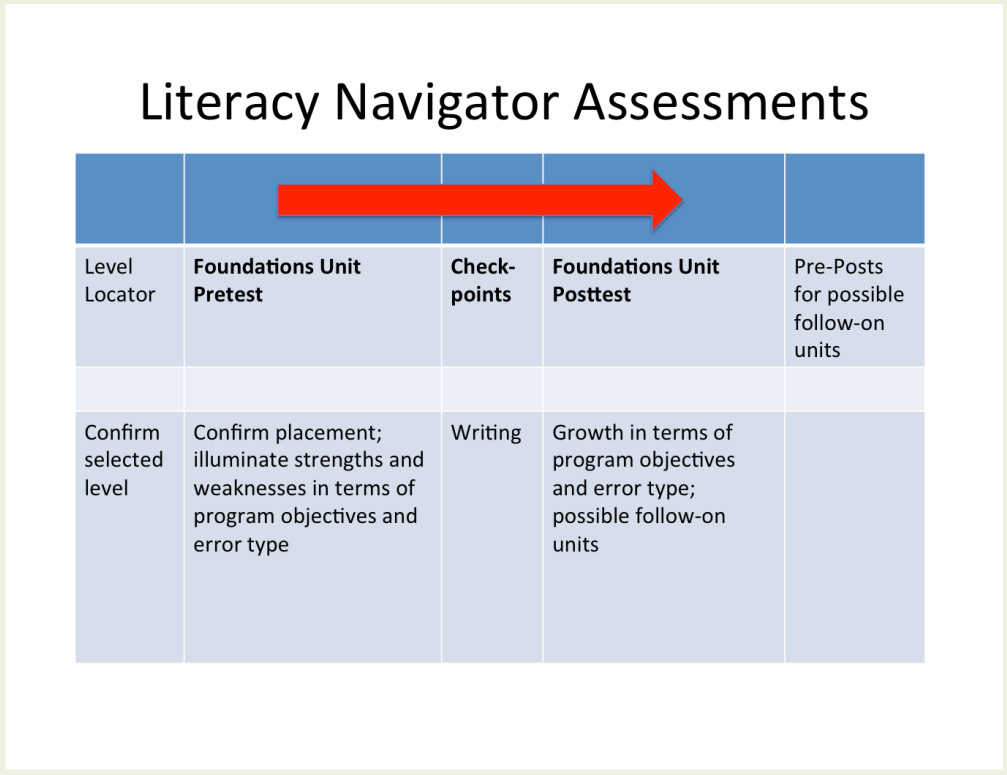
Literacy Navigator

The Program

Literacy Navigator is also an intervention program, designed for students who are having trouble keeping up with their regular classroom instruction and need additional focused teaching around informational text comprehension. It consists of a foundation module and several follow-on modules, each providing instruction in comprehension of informational text.

The assessments for Literacy Navigator (Figure 7) are also very carefully designed, and the reports feature diagnostic information similar to those just described for Mathematics Navigator. The roster reports are organized just as they are for Mathematics Navigator. They provide a listing of what each student gave as an answer for each question, and a hyperlink to the question itself so that the teacher can view the question and the option choices. The texts used are not provided on line; teachers must refer back to the actual tests themselves to view the text, but the actual items are viewable through the hyperlinks.

Figure 7: Literacy Navigator assessments



These roster reports also show the percent of students answering each item correctly, and the percent choosing each option. Wrong option choices are shaded yellow. In addition, any item where a large number of students chose the same wrong answer is shown so that teachers can focus on whole class misunderstandings.

The assessments and reports follow the pattern established for Mathematics Navigator, except that instead of a grade level screener there is a test to confirm the appropriateness of the grade level chosen for a particular group of students. Figure 8 shows the reports provided for Literacy Navigator. Please note the similarities to the structure for Mathematics Navigator.

In Literacy Navigator, there are program objectives sub-scores shown on the roster report as well as total scores. The test is broader than the Mathematics Navigator tests, where the total score relates to only one specific strand of mathematics. The use of sub score information gives finer grained information than a total comprehension score. Information is given about student's ability to accurately retrieve details, make inferences, link information, deal with issues of pronoun reference, handle mid-level structures such as cause and effect, sequence and problem/solution, and word study concepts.

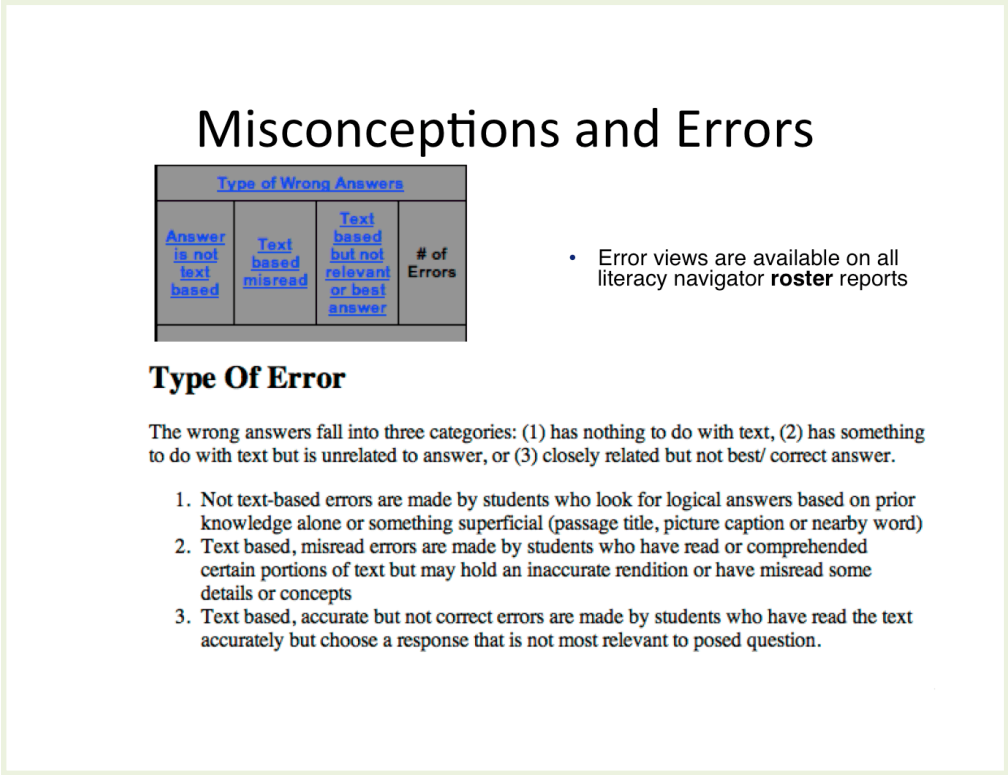
The primary diagnostic information comes from an analysis of error patterns. This is like the misconception analysis for mathematics. Each option choice is coded as being either a non-text- based response, a text-based misread or a text-based response that is accurate but not the right answer to the question posed. The percent of wrong answers falling into each of these categories is then reported for each student to show the kind of error being made. This is extremely useful information to a teacher. Two students with the same number of errors, but for one the errors are all non-text-based and the others are text-based but just not the right answer pose two different challenges for instruction. Figure 9 shows this information.

Figure 8: Literacy Navigator reports

Literacy Navigator Reports

	Level of Analysis	Instructional Purpose/ Value Added
Level Locator Reports: Summary and Detail	Class, Grade, School, District, AYP subgroup	Program level needed by student, item level data, students who need lower/ higher levels of program
Roster Report: Per Unit	Class, Grade, School, District, AYP subgroup	Total scores by student, percent correct by item and for each distractor, scores on program objectives, grouping by types of wrong answers
Checkpoint Report: Per Unit	Class, Grade, School, District, AYP subgroup, Achievement Group	Quick snapshot of whether students are on target to-date and are making sufficient progress
Aggregate Portfolio Report: high level snapshot across units	Class, Grade, School, District, AYP subgroup, Achievement Group	# of students tested, total scores, growth report (scaled scores), comparison with group
Class Profile Report	Class	Report that summarizes number of times that teacher observes particular skills
Student Achievement Report: Per Unit	Student	Total score, growth report (scaled score), comparison with group, item level data, likelihood of having specific error patterns
Student Portfolio Report: high level snapshot of all units and performance	Student	Total scores, benchmarks for performance and goal setting, progress indicators, growth report, assessment of meeting goals, comparison with group

Figure 9: Literacy Navigator misconceptions and errors



Design Issues

From a design perspective, there are at least four opportunities for a student to choose an option that falls into one of the three kinds of errors. This allows for stable estimation of the pattern of errors a given student is making in response to a specific level of complexity of text.

One of the most important ways that these designs came about was the result of developing the assessments as the curricula for both mathematics and literacy programs were being developed. Working alongside of the curriculum developers allowed for the alignment of the assessment with the intention of the curriculum designers and the allowed for the capturing of the diagnostic approaches within the curricula themselves. Thus, what emerged was a very carefully designed and aligned approach that allowed the reports to follow the design of the curriculum and the assessments in a way that makes them maximally useful to teachers as they proceed with instruction.

Summary

The diagnostic use of curriculum embedded assessments is an important ingredient in a successful formative assessment program. The fundamental design principles that these assessments illustrate relates primarily to the issue of validity, as discussed in some length in the lead article in this issue. If test is to be valid to serve classroom teachers, it must be designed with a carefully planned set of reports that will address their needs. Teachers need the assessments to be

helpful to them in planning differentiated learning, finding the strengths and weaknesses that their students have so they can be addressed on an individual pupil basis.

It is in the design elements of these reports that will make or break the use for which the information is intended. Having reports that show individual student misconceptions or error patterns is the key ingredient of the reports described here, and they are critical to the teacher's ability to group students appropriately for instruction, to address identified needs, and to tailor additional formative daily assessment activities to reflect the underlying misconceptions or pattern of responses that students are displaying.

In addition to the misconception and error patterns, the design of the reports allows teachers to have a bird's eye view of the whole class performance, by providing the overall item analysis information with hyperlinks that allow teachers to view items as they are examining how the whole class performed. Highlighting any places where many students chose the same wrong answer, and viewing the item with its option choices in a direct and immediate manner allows he teacher to view larger chunks of performance gaps that can be addressed.

The tests, obviously, must be carefully designed to allow for the generation of the reports that support valid inferences about student behavior that gets reflected in the reports. Selecting wrong answer choices in the preplanned way that both the mathematics and literacy assessments were done allows the teacher to see first, if the students are demonstrating reliable error patterns, and second, to have those reliable patterns reported on in a way that allows for customizing classroom practice.

About the Author

Betsy Taleporos has recently retired as the Director of Assessment for America's Choice. She managed all the research, evaluation, and assessment work for this organization which has had a major impact in the Standards-based education movement and in national School Reform efforts. She was responsible for the development of approximately 300 mathematics and literacy assessments, including performance based and multiple choice formatted tests, all of which are curriculum embedded and are directly linked to classroom practice. Prior to joining America's Choice, she managed large-scale test development projects in English Language Arts and Mathematics for several major national test publishers. Before that, she directed the assessment efforts in New York City managing the efforts in test development, psychometrics, research, analysis, administration, scoring, reporting and dissemination of information. In that capacity she also served as the New York City site coordinator for the New Standards project. Betsy brings a strong background and expertise in areas of practical application, aligning instruction and standards and assessments, and in academic research and teaching at the graduate and undergraduate level for New York University, Adelphi University and Long Island University.

Taleporos, E. (2012) Periodic Assessments and Diagnostic Reports. *Educational Designer,* 2(5).
Retrieved from: http://www.educationaldesigner.org/ed/volume2/issue5/article17/