PEARSON

assessment report

The Distractor Rationale Taxonomy: Enhancing Multiple-Choice Items in Reading and Mathematics

Kelly V. King

Doug A. Gardner

Sasha Zucker

Margaret A. Jorgensen, Ph.D.

July 2004

The Distractor Rationale Taxonomy: Enhancing Multiple-Choice Items in Reading and Mathematics

Introduction

Recent education reform legislation, especially the *No Child Left Behind Act* of 2001 (NCLB), has highlighted discussion concerning the relationship between assessment and classroom instruction. The expectation is that if assessment results can be used to improve classroom instruction, student performance on accountability measures will also improve. Concurrent with this expectation is growing support for assessments specifically designed to improve classroom instruction (The Commission on Instructionally Supportive Assessment, 2001; National Research Council, 2003). To accomplish this goal, an assessment must provide educators with information about what a student knows and what additional instruction or intervention the student requires to likely attain desired learning outcomes (Linn and Gronlund, 2000; Nitko, 2004).

The time is right to explore new methods of item development that tighten the relationship between the way items are written and the way teachers teach. A traditional multiple-choice item provides little information about a student's level of understanding in a content standard beyond whether he or she has selected the correct response. However, multiple-choice items written using the *distractor rationale taxonomy*—developed by Pearson Inc. (Pearson)—reveal a student's breakdown in understanding through his or her incorrect answers. An assessment system that incorporates this methodology, such as *Stanford Learning First,*™ can indicate a student's instructional needs in a subject area and thereby contribute to the development of a focused intervention plan. This paper discusses the distractor rationale taxonomy's background and its valuable practical application in the construction of multiple-choice items.

Multiple-choice Items

The multiple-choice item format is widely recognized as the most versatile and efficient item format for standardized tests in a wide range of academic subject areas (Airasian, 2001; Kubiszyn and Borich, 1987; Linn and Gronlund, 2000; Nitko, 2004). The uses of multiple-choice items and the rules for authoring them are well documented in publicly available literature and in the proprietary training materials of test publishers. Because only a count of the correct responses is required to obtain a student's score, assessments that use multiple-choice items produce highly objective results (Kubiszyn and Borich, 1987). As such, they have a long history of helping instructors and policy makers understand what students know and can do.

A multiple-choice item consists of one or more introductory sentences, called the *stem*, followed by a list of two or more possible answers. Typically, only one answer, referred to as the *key*, is correct. The incorrect answers are referred to as the *distractors* and typically represent common errors that are equally plausible only to students who have not attained the level of knowledge or understanding necessary to recognize the key as the correct answer (Linn and Gronlund, 2000; Popham 2000; Nitko, 2004). An example of a multiple-choice item follows:

What is the main idea of the passage "Frogs and Toads"?		[Stem]
A	**Frogs and toads share many differences and similarities. ***	[Key]
B	**Frogs live closer to water than toads.**	[Distractor]
C	**Frogs and toads are like cousins.**	[Distractor]
D	**Frogs are different than toads.**	[Distractor]

Rethinking Distractors

For decades, Pearson has followed best practice in the development of multiple-choice items. In particular, an item's distractors have been designed primarily to draw uninformed or unskilled students away from the correct answer (Popham, 2000; Nitko, 2004). A high-quality item should have distractors that represent equally plausible, common errors that a student might make. During development of an assessment, statistical analysis can ensure each distractor is equally plausible. If any distractor is not selected with the frequency of the others in the same item, it detracts from the overall quality of the item and must be either eliminated or revised (Linn and Gronlund, 2000). Hence, the writing of distractors has been guided primarily by editorial and stylistic concerns and has received less investigation than other elements of multiple-choice items.

Recently, authorities in educational assessment have suggested extending the functional role of distractors to include a new purpose: identifying the nature of a student's misunderstanding. Nitko (2004) observes that "the distractor a student chooses may give you diagnostic insight into the difficulties the student is experiencing" (p. 142). Popham (2000) also recognizes the potential of distractors to represent "the categories of incorrect responses that students make" and thereby allow teachers to "follow up with additional instruction based on the most common sorts of errors made by an individual student or a group of students" (p. 244).

Educational researchers in the state of Washington have performed studies that support a new function for distractors. In an approach significantly different from Pearson's, their experimental methodology encodes an exhaustive list of common student misconceptions (or, to use the researchers' preferred terminology, *facets of understanding*) in physics topics. This list is used to interpret the incorrect responses of students to items of various formats (Minstrell, 2000; National Research Council, 2003). Combined with other information gathered during the assessment, such as interviews, the student's incorrect response corresponds to a relevant learning activity (Minstrell, 2000).

Clearly, the patterns in a student's incorrect responses to items can provide powerful, valuable information to guide instruction. However, such guidance can only be obtained from a multiple-choice item when its distractors have been developed in a systematic and meaningful way. Pearson's methodology builds upon this discussion and research and makes the leap forward that brings this innovative capability to multiple-choice items in formative assessments.

Pearson's Distractor Rationale Taxonomy

Pearson has developed a system for using distractors to collect and report data that reveals a student's level of understanding of a subject area.[1] Pearson's long-established quality guidelines for writing items dictate that each distractor for an item must include a *rationale*—an explanation of why that distractor is the wrong response. When organized into a taxonomy that reflects levels of understanding, distractor rationales enable the development of multiple-choice items that yield a rich breadth and depth of information about why students make incorrect choices (Gardner, 2004; King, 2004).

Pearson's distractor rationale taxonomy indicates four levels of understanding: level one includes the most fundamental errors; levels two and three correspond to responses that, while incorrect, indicate increasing sophistication in the student's response; and level four represents the correct response (an item's key). The taxonomy can be applied to many different subject areas including reading, English language arts, and mathematics. Using the distractor rationale taxonomy, item writers can design distractors to contribute to an analysis of a student's pattern of misunderstanding in a subject area.

[1] Research is ongoing to further document the efficacy of this approach and the extent to which it improves instruction, student learning, and student performance on accountability tests. Pearson will report this research as it is completed.

Reading Item Examples

Table 1 (from King, 2004) lists and describes the types of errors in reading that correspond to each of the four levels of understanding in the distractor rationale taxonomy:

Table 1. A distractor rationale taxonomy for reading items.

Level of Understanding	Student Error
Level 1	Makes errors that reflect focus on decoding and retrieving facts or details that are not necessarily related to the text or item. Student invokes prior knowledge related to the general topic of the passage, but response is not text-based. These errors indicate that the student is grabbing bits and pieces of the text as he or she understands them, but the pieces are unrelated to the information required by the question being asked.
Level 2	Makes errors that reflect initial understanding of facts or details in the text, but inability to relate them to each other or apply them to come to even a weak conclusion or inference. The student may be focusing on literal aspects of a text or on superficial connections to arrive at a response.
Level 3	Makes errors that reflect analysis and interpretation, but conclusions or inferences arrived at are secondary or weaker than ones required for correct response. A distractor may be related to the correct response in meaning, but be too narrow or broad given the circumstances.
Level 4	Correct response.

The following two examples demonstrate how main idea and vocabulary-in-context items can be written to use the distractor rationale taxonomy. These items are associated with a grade 3 reading passage titled “Frogs and Toads,” which is not reproduced here.

What is the main idea of the passage "Frogs and Toads"?

A Frogs and toads are cute.
[Level 1: prior knowledge, not text-based]

B Toads have shorter legs than frogs have.
[Level 2: text-based detail unrelated to main idea]

C Frogs are different than toads.
[Level 3: only part of main idea]

D Frogs and toads share many differences and similarities. *
[Level 4: correct response]

Note that each level of understanding is represented in the item's options. Ideally, an item will include distractors that correspond to each distinct level of understanding in the taxonomy. In some items that assess very simple content standards, such as the recall of specific facts, full representation of each level may not be possible. However, an item capable of representing each level in its distractors provides the best assessment of what a student knows and can do relative to a content standard (King, 2004).

Compare the previous design to an item with the same stem and key, but with distractors that were developed only for their traditional function.

What is the main idea of the passage "Frogs and Toads"?

A Frogs live closer to water than toads.

B Frogs and toads are like cousins.

C Frogs are different than toads.

D Frogs and toads share many differences and similarities. *

All distractors are essentially level 3: Each is related to the main idea but is not the *best* answer.

In the second version of the item, the distractors are effective in their traditional role of presenting alternatives that would be plausible to a student without the sufficient knowledge to recognize the key. However, the distractors do not provide the information necessary to determine how close a student is to meeting the content standard. The only information the item provides is whether the student was successful in selecting the key. Although the educator can be certain that the student has selected an incorrect response, it is difficult to determine *why* the student selected that response. An item developed using the distractor rationale taxonomy clearly provides this missing piece of the puzzle. Instead of simply identifying whether a student knows a concept or not, this taxonomy delves into the cognitive process of the student to identify a possible misconception or misunderstanding.

The next example tests vocabulary-in-context using the same reading passage.

Read this sentence from the passage "Frogs and Toads."

> Both frogs and toads have a tail at first that <u>disappears</u> when they get older.

Which word has the same meaning as <u>disappears</u> as it is used in this sentence?

A	**disagrees**	[Level 1: look-alike word]
B	**vanishes ***	[Level 4: correct response]
C	**can be seen**	[Level 2: antonym]
D	**becomes small**	[Level 3: related to the meaning, but not precise]

This vocabulary-in-context item demonstrates another application of the taxonomy. Because understanding unfamiliar words in context requires students to make inferences, the taxonomy allows the item to have distractors spread across all levels of understanding, thereby helping the teacher identify the student's possible misunderstanding. As shown in the next example, the same item structured in the traditional manner does not provide this insight into the student's breakdown in understanding if he or she chooses an incorrect option.

Read this sentence from the passage "Frogs and Toads."

Both frogs and toads have a tail at first that <u>disappears</u> when they get older.

Which word has the same meaning as <u>disappears</u> as it is used in this sentence?

A turns green

B vanishes *

C jumps

D breathes

All distractors are related to frogs, not to the meaning of the word "disappears."

This traditional item structure helps the teacher understand whether the student can apply the skill of understanding unfamiliar words in context, but cannot give any further insight into the student's level of understanding when applying the skill.

Mathematics Item Examples

Table 2 (from Gardner, 2004) lists and describes the types of mathematics errors that correspond to each of the four levels of understanding in the distractor rationale taxonomy:

Table 2. A distractor rationale taxonomy for mathematics items.

Level of Understanding	Student Error
Level 1	Makes errors that reflect combinations of conceptual misunderstanding and unfamiliarity with how to relate operational procedures to problem contexts. Student attempts to implement strategies that are unrelated to the task at hand. These errors may indicate that the student has an inordinate dependence on information that is explicitly stated in item stimuli, and is lacking the sophistication to correctly answer the question being asked.
Level 2	Makes errors that reflect some sophistication and computational ability, but that demonstrate an inadequate conceptual framework and flawed reasoning in support of conclusions or inferences.
Level 3	Makes errors that reflect definite levels of sophistication in analysis and conceptual knowledge, but that are flawed by inconsistent reasoning or computational weakness.
Level 4	Correct response.

This taxonomy provides mathematics item writers with a clear way of classifying each distractor so that it explains why a student has made an incorrect choice. The following example (from Gardner, 2004) demonstrates the use of the distractor rationale taxonomy in the design of a number and operation item.

Janice spent $222.46 before tax to buy a climbing rope that sold for $3.50 per meter. What is the greatest number of meters of the rope she could buy at that rate?

A	**7786.1 m**	[Level 1: incorrect operation with place-value error]
B	**778.61 m**	[Level 2: incorrect operation, correctly applied]
C	**635.6 m**	[Level 3: correct operation with place-value error]
D	**63.56 m ***	[Level 4: correct response]

As in the reading items, the use of the distractor rationale taxonomy in the design of mathematics items provides information about the nature of the student's error. Compare the previous item to the following item which has the same stem, but different distractors:

Janice spent $222.46 before tax to buy a climbing rope that sold for $3.50 per meter. What is the greatest number of meters of the rope she could buy at that rate?

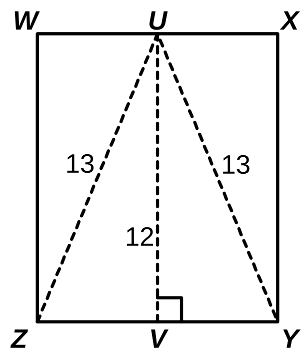
A **0.6356 m**
B **6.356 m**
C **63.56 m** *
D **635.6 m**

All distractors are essentially level 3. Each is based on a computation error related to place value.

The distractors in the second example each represent an error at the same level of understanding and do not provide any further information about the student's ability to perform tasks requiring this mathematics skill.

The distractor rationale taxonomy can also be applied to measurement items, as seen in the following example (from Gardner, 2004):

In the diagram showing rectangle *WXYZ*, the lengths of $\overline{ZU}$, $\overline{VU}$, and $\overline{YU}$ are labeled in meters.



What is the perimeter (in meters) of rectangle *WXYZ*?

A 34 [Level 3: student correctly applied Pythagorean theorem, but failed to double length to account for full length of side]

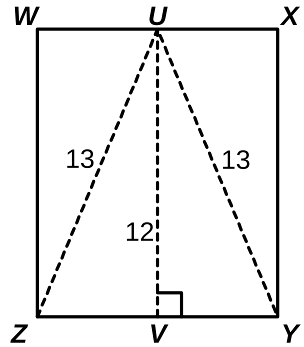
B 38 [Level 1: student simply added the values apparent in the diagram]

C 44 [Level 4: correct response]

D 120 [Level 2: applied Pythagorean theorem, but confused concept of area with concept of perimeter]

Compare this item, which uses the distractor rationale taxonomy, to an item with the same stem but with distractors designed for their traditional function:

In the diagram showing rectangle *WXYZ*, the lengths of $\overline{ZU}$, $\overline{VU}$, and $\overline{YU}$ are labeled in meters.



What is the perimeter (in meters) of rectangle *WXYZ*?

A 44 *
B 50
C 120
D 156

Options are based on unwarranted assumptions about the implied lengths of the rectangle's sides as well as the meaning of perimeter. The option set is visually balanced.

Again, the distractors in this item serve their traditional purpose, but they do not provide the insight that can be obtained from distractors that are designed using the distractor rationale taxonomy.

Conclusion

In products that use the distractor rationale taxonomy, multiple-choice items can reveal patterns in a student's level of understanding in a content standard. *Stanford Learning First,* Pearson's forthcoming assessment system, incorporates the distractor rationale taxonomy in the items of *Classlinks,*™ its formative assessment component. This feature of *Stanford Learning First* guides teachers toward instruction that specifically and directly addresses a student's breakdown in understanding in a content standard. To maximize the benefits of instructional intervention, teachers can group students based on their level of understanding. By linking these patterns to instructional activities, intervention plans, and professional development for educators, *Stanford Learning First* can immediately contribute to more effective classroom instruction. Pearson's distractor rationale taxonomy represents a significant innovation in the testing industry.

References

Airasian, P. W. (2001). *Classroom assessment* (4th ed.). Boston: McGraw-Hill.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

The Commission on Instructionally Supportive Assessment. (2001, October). *Building tests to support instruction and accountability: A guide for policymakers.* Retrieved from http://www.nea.org/accountability/buildingtests.html on April 16, 2004.

Gardner, D. A. (2004). *Distractor rationale taxonomy for mathematics.* Unpublished training materials, Pearson Inc.

King, K. V. (2004). *Distractor rationale taxonomy for reading.* Unpublished training materials, Pearson Inc.

Kubiszyn, T., & Borich, G. (1987). *Educational testing and measurement*: *Classroom application and practice* (2nd ed.). Glenview, IL: Scott, Foresman, and Company.

Linn R. L., & Gronlund N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice Hall, Inc.

Minstrell, J. (2000). Student thinking and related assessment: Creating a facet-based learning environment. In Raju, N. S., Pellegrino, J. W., Bertenthal, M. W., Mitchell, K. J., & Jones, L. R. (Eds.) *Grading the nation's report card: Research from the evaluation of NAEP* (pp. 44–73). Washington, DC: The National Academies Press.

National Research Council. (2003). *Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment. Workshop report.* Washington, DC: The National Academies Press.

Nitko, A. J. (2004). *Educational assessments of students* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Popham, W. J. (2000). Educational measurement (3rd ed.). Boston, MA: Allyn and Bacon.

Additional copies of this and related documents are available from:
Pearson Inc.
19500 Bulverde Road
San Antonio, TX 78259
1-800-211-8378
1-877-576-1816 (fax)
http://www.pearsonassess.com