The Role of Item Response Attractors In Modified Items for Students with Disabilities

April 15, 2009

Michael C. Rodriguez

University of Minnesota

Peter Beddow, Elizabeth Compton, Stephen Elliott, Ryan Kettler, Andrew Roach

CAAVES Team

Paper presented at the annual meeting of the

National Council on Measurement in Education

Symposium on Designing and Evaluating Modified Items for Students with Disabilities:

Research Results; S. N. Elliot, Organizer

The Role of Item Response Attractors
In Modified Items for Students with Disabilities

The activity of item writing has recently received a boost in attention, largely because of the high-stakes demands that have been placed on achievement testing through NCLB requirements, the increasing attention to international assessments (e.g., PISA, TIMSS), and even a renewed interest in formative classroom assessment. Within this attention has been a focused effort to create test items that are accessible to students with a wide range of abilities, particularly those in the NCLB 2% category (more significant cognitive impairments). A large part of the attention also is due to the introduction of the principles of Universal Design into the item-development process and attention to language complexity (that is construct irrelevant) providing greater access to students learning English as a second language.

Item writing has received limited attention in the measurement research literature, where most argue that the science of item writing is underdeveloped. Some of this work began in the 1920s, where testing specialists conducted research on variations of item formats. Only a handful of the item-writing guidelines in the current literature have been tested empirically, including the number of options used, the use of None-of-the-above and All-of-the-above options, the complex type-K format, the use of negatively phrased stems, and option length. Issues related the number of options have been studied about three times as much as any of the other issues.

Attention to distractors has been the most popular area for research on item writing. Some of the earliest leadership on item writing was provided by Ebel (1951) in his seminal chapter on the topic. Haladyna and Downing (1988) presented a framework for developing functional distractors. Haladyna and Downing later presented a taxonomy of item writing guidelines based on a review of textbook author recommendations (1989a) and supplemented

that with validity-related evidence from the empirical research literature (1989b). Haladyna (2004) provides contemporary leadership in the area of item writing with his text *Developing and Validating Multiple-Choice Test Items*, currently in its 3rd edition. Regarding the number of options for a multiple-choice item, the prevailing guideline is to develop as many functional distractors as possible (Haladyna, Downing, and Rodriguez, 2002); however, empirical research suggests that three is adequate and perhaps optimal (Rodriguez, 2005).

From Distractors to Attractors

One aspect of combining concepts of Universal Design (CUD, 2008), good item-writing principles (Haladyna, Downing, & Rodriguez, 2002), and cognitive load theory (Clark, Nguyen, & Sweller, 2006), suggests the creation of three-option items with careful attention to the item response attractors, traditionally referred to as distractors. The use of the term distractors (or "foils") stems from classical item writing literature (Ebel, 1951). The intent is that the incorrect options "distract" the student with limited knowledge and understanding. In the context of the principles described above (which will be described more fully during the presentation), a more productive intent of an incorrect option is to attract those students with a specific misconception or error in knowledge and thinking.

This relabeling of the incorrect option then explicitly requires greater attention to the design of the "attractors" because they must then contain information about misconceptions or errors in order to attract the right students. This is really not much different than what we would find in item-writing guidelines, but in practice, most items appear to be written such that incorrect options are not functioning well, largely because they do not conform to these principles (Rodriguez, 2005).

The Role of Attractors

Attractors play the most important role in a multiple-choice item. They present the challenging aspects of an item. We find the difficulty of an item most easily manipulated by the nature of the attractors, particularly in their proximity. Item writing researchers have found that the plausibility and proximity (similarity) of the attractors has a much greater impact on item difficulty than do characteristics of the stem (for example, whether the stem is a complete question or open-ended statements completed by the options) (Ascalon, Meyers, Davis, & Smits, 2007; Haladyna, Downing, & Rodriguez, 2002; Smith & Smith, 1988). Consider the following example:

1. Who was elected President of the United States in 1932 during the Great Depression?

   A. Daniel Boone
   B. Dwight D. Eisenhower
   C. Ronald W. Reagan
   D. Franklin D. Roosevelt

2. Who was elected President of the United States in 1932 during the Great Depression?

   A. Calvin Coolidge
   B. Herbert Hoover
   C. Franklin Roosevelt
   D. Theodore Roosevelt

First, you should note that the options are in alphabetical order. However, in the first item, all options are not plausible, since Daniel Boone was not a president; remaining options are quite different in plausibility, since Reagan was president far more recently than Eisenhower or Roosevelt. In the second version of the item, all options were president, three of which were president during 1923 and 1945. Theodore Roosevelt might be an interesting option since he had the same last name as the correct response; however, this might present a clue since there are two Roosevelt's – "It must be one of them." Nevertheless, the first version of the item is likely to be much easier than the second. Perhaps the most important consideration is that this is not a

particularly interesting item since it is tapping simple recall. Consider the following item (based on a similar item by Haladyna (1999):

3. What is the most effective method to reduce the internal air temperature of a house in a humid subtropical climate?

   A. Fan
   B. Evaporative cooler
   C. Air conditioner
   D. Dehumidifier

4. What is the most effective way to cool a home in a humid climate?

   A. Air conditioner
   B. Evaporative cooler
   C. Fan

Item number 3 presents several problems. First, the stem is wordy and unnecessarily technical. One of the options, “Dehumidifier”, contains part of a key word in the stem, “humid”, which presents a “clang association” leading to errors that are construct irrelevant. Depending on the region, the use of an “evaporative cooler” may or may not be familiar (this is also a curriculum issue) – the simple evaporation of water tends to cool the air. So the question is a complex one, requiring students to evaluate each option, and select the most effective, since more than one will cool the air (humidifiers reduce the humidity, but tend to create more heat in the process). A minor point is that the options are not in any particular order, where alphabetical, numerical, or some other logical order provides for a standard method of ordering options. Item number 4 corrects most of the faults of the first version and retains the complex nature of the question.

When considering modifications of items for students with cognitive impairments, the application of strong item-writing guidelines will improve the quality of resulting responses and improve measurement overall – but also for all students. Here, I focus on the role of the attractor.

From the above examples, it is clear that the options are an important part of the item and the functioning of the attractors is essential. With the exception of classroom assessment, it is difficult to defend the use of humorous or blatantly absurd options – which research suggests should only be used if they are consistent with the instructional and personal characteristics of the instructor (McMorris, Boothroyd, & Pietrangelo, 1997). The presence of obviously irrelevant options takes time (an important resource) for students to consider and reason about its relevance (which may be a more difficult task for students with cognitive impairments), potentially distracts students from thinking clearly about the construct, reduces the measurement power of the item, and eliminates the chance that additional information about the presence of misconceptions or reasoning errors held by students.

A nonfunctioning option is one that is not selected by students or does not discriminate between high and low ability students. Nonfunctioning options typically are options that are not plausible. The elimination of nonfunctioning options then promote several goals in making test items more accessible to all students, particularly by reducing the per-item testing time and eliminating potential sources of confusion. By using the label “distractor”, these issues are not central concerns to the item writer – it may not be important as to why an option distracts a student. When we enter the item writing task or item modification task using the language of “attractors”, our attention is focused on the explicit role of the attractor in presenting a plausible challenge to the correct response. At the same time, we need to make sure that the attractors are attracting the right students, those students with misconceptions or reasoning errors, those students who tend to be of relatively lower ability.

Current Study Purpose

The current study originated during an IES funded project, The Consortium for Alternate Assessment Validity and Experimental Studies (CAAVES), a 4-state collaboration to modify

state assessments. Two manuscripts have been produced that summarize findings and the work in the project, including Kettler et al. (under review) and Elliott et al. (in press). Alternate assessments of modified achievement standards are allowed by the US Department of Education to provide access to the assessment of state standards for students receiving special education services and who are unable to attain proficiency on the regular assessment. Through this project, researchers and states have modified test items to make them more accessible and reducing cognitive load, improving the validity of results for those students who otherwise would not be able to display their knowledge and skills.

In the CAAVES project, an experimental test design study including students with and without disabilities, where 4-option test items from an existing data bank containing item statistics were modified by removing one option that was either not functioning (based on item statistics) or was implausible (i.e., did not contain useful information about misconceptions, based on logical analysis of the options). Additional modifications were made to reduce the complexity of language and sentence structure and provide greater access to the intent of the item, including the use of carefully selected graphics or pictures. The study included a carefully balanced design with rotation of modified items from the beginning to the end of the test across students, such that each student was exposed to items in original and modified format. The results of this study have been analyzed in terms of shifts in item difficulty and test score reliability.

Method

Participants

The sample included 755 eight grade students from four states, including students with disabilities (n=486) and without (n=269). Several elements of the original design were ignored for this purpose, including the difference between students with disabilities that were eligible for

the alternate assessment (*n*=250) and those with disabilities that were not eligible (*n*=236). The students were approximately 58% male and 69% White.

Measures

The CAAVES study included mathematics and reading tests, both composed of 39 multiple-choice items. The items were provided by Discovery Education Assessment (DEA) from a pool of items identified to meet common state standards in both areas. The mathematics test included 20 numbers items requiring decoding of mathematical symbols and basic operations and 19 data items requiring basic arithmetic operations. The reading test included 20 comprehension items and 19 vocabulary items. Each item was modified using the principles discussed above and those summarized in the Test Accessibility and Modification Inventory (TAMI, Beddow, Kettler, & Elliott, 2008). The TAMI includes guidance for developing (modifying) the passage or item stimulus, item stems, visuals, answer choices, page format and layout, and issues related to fairness.

When the tests were assembled, they contained three sets of 13 items (39 items in total), where one set was in original format, one set was modified, and the third set was modified with the addition of reading support in the form of a recorded voice that read item directions and stems). The three forms of items were rotated across each of the three sets of 13 items and across the three positions of the test (first, second, and third set of 13 items). This balanced item order and item format. The tests were administered by computer. Coefficient alpha was .89 for reading and .85 for mathematics.

Study Design

A quasi-experimental design was employed with three groups of students: students without disabilities and students with disabilities that were either eligible or not for the alternate

assessment (determined through specific participation criteria). Each student completed one set of 13 items in each of the three formats: original and modified with and without reading support.

Previous research (Elliott et al., in review) has thoroughly studied the effect of condition and group on test score reliability and item difficulty and discrimination. Although there were significant group by condition interaction effects on test score reliability, all differences were not very meaningful, where all differences were less than .06. No matter how else the data were cut (group, condition, order of item set within the test form), all other differences were less than .02. The researchers suggested that this provided evidence to support systematic modifications without undermining score consistency.

Elliott et al. (under review) then used the Rasch model to assess modification effects on item difficulty (Rasch was used to equate group ability distributions). They found item difficulties to be reduced through item modification on both tests and this effect was greater for eligible students than for students without disabilities. This supported the researchers' "interaction hypothesis", such that eligible students experience a greater benefit from the modification than students who are not eligible to participate in the alternate assessment.

Current Analyses

To continue the analyses of these data, I examined the functioning of the options for each item. All analyses were across students (student group was not of interest for these purposes) and the reading support aspect was ignored (item format was considered to be either original or modified). First, I reexamined the effect of modification on overall item statistics, including item difficulty and discrimination (classical test statistics). Then, the attractor-discrimination values were examined for each item. The attractor discrimination is the point-biserial correlation between the selection of an attractor and the total score. Ideally, attractor discrimination indices should be negative, indicating that the selection of an attractor is associated with a lower total

score overall. Item difficulty and discrimination and the discrimination index for each option were estimated using Winsteps 3.65, a Rasch analysis program, which estimates classical test statistic values.

With modification, we should see an increase in the item *p*-value (more students respond correctly making the item easier) and since the measurement properties should improve, we should observe this in an increase in the item-discrimination. Finally, the two attractors remaining should result in stronger discrimination as well, where the discrimination index is negative and larger in the modified version than the original version.

Recall that the modifications consisted of up to three conditions, including the removal of one option, simplification of language complexity, the possible addition of a graphic, and possible reorganization of the item layout. So, although comparison are made to the synthesis results for changing the number of options, these results are based on a package of modifications, not just the removal of one option.

Results

Item Difficulty

Classical item difficulty was examined in terms of the item *p*-value (proportion correct); here I present the percent correct for ease. Rodriguez (2005) reported an average increase in the percent correct of about 4.4% when reducing the number of options from 4 to 3 across 36 studies in his review. Here, the average change in item difficulty was about 6% for mathematics and 10% for reading, making the items easier on average for both tests. Overall, 6 of 39 items became more difficult in mathematics; only two items became more difficult in reading; the remaining became easier. Table 1 contains summary statistics for mean item difficulty changes between original and modified formats.

Table 1

Summary Statistics of Differences in Item Percent Correct from Original to Modified Format

	Minimum	Maximum	Mean	SD
Mathematics	-7%	18%	6%	6%
Reading	-12%	35%	10%	10%

Item Discrimination

Item discrimination was based on the corrected point-biserial correlation between the item and the total score (excluding that item)[1]. Rodriguez (2005) found an average increase in item discrimination of .03 when the number of options was reduced from 4 to 3 based on 30 studies. Here, the average change in item discrimination was -.05 for mathematics and -.01 for reading, both in the unexpected direction. Overall, in mathematics, 11 of 39 items increased in discrimination; in reading, 17 items increased in discrimination. Table 2 contains summary statistics for mean item discrimination changes between original and modified formats.

Table 2

Summary Statistics of Differences in Item Discrimination from Original to Modified Format

	Minimum	Maximum	Mean	SD
Mathematics	-.30	.16	-.05	.08
Reading	-.16	.32	-.01	.11

Attractor Functioning

[1] Item discrimination results were transformed to Fisher's Z for statistical analyses to account for the skewed nature of the sampling distributions of correlations.

An effective attractor is one that attracts students with misconceptions or errors in thinking and reasoning. There are two common indices that help us assess the effectiveness of an attractor, including the response rate (number of students selecting the attractor) and the attractor discrimination (corrected point-biserial correlation). Table 3 contains summary statistics for changes in attractor functioning from original to modified format. For both mathematics and reading, on average, attractors became more discriminating in modified format. In mathematics, 71% of all attractors across the 39 items became more discriminating in modified format; in reading, 78% of all attractors became more discriminating.

Table 3

Summary Statistics of Differences in Attractor Discrimination from Original to Modified Format

	Minimum	Maximum	Mean	SD
Mathematics	-.27	.20	.04	.08
Reading	-.23	.23	.06	.09

A complete item-by-item analysis of attractor functioning is provided in the Appendix. For each item, the response frequency to every option and the option point-biserial correlations are listed, for the item in original and modified format.

In addition to examining the change in item and attractor statistics, I examined the response rates for the attractors as well as the change in response rate for each attractor. To facilitate the comparison, I computed the percent responding to each attractor, using only the three options that remained in the modified item, but for both the original and modified items – thus, I ignored the responses to the option that was deleted and recomputed the percent responding based on the *n* responding to the remaining options. There were two reasons for

considering an analysis of attractor response rates. First, If this attractor was removed, the remaining attractors should be plausible and differences in response rates between attractors should be reduced – attractors should be plausible and relatively equally attractive.

These results are less conclusive and should be interpreted with more caution. Two limitations are significant. In the original version of the item, if an option was eliminated, it is not clear that the modified option response rates would be observed because modifications were also made to the language and layout of many items. The recomputation of option response rates on original items is a simple comparison, ignoring the fact that those students would have selected other options (possibly the correct option) if that eliminate option were not present. In addition, when an option is removed, the item tends to become easier, meaning that more students are selecting the correct response and not one of the attractors. Nonetheless, the examination of attractor response rates is useful to complete this analysis of the role of attractors.

On the mathematics test, there is an overall increase in the response rates to attractors from the original to modified formats of items. On average, attractor response rates decreased 0.3% (meaning, 0.3% fewer students selected the remaining attractors in the modified version than the original version), when examining only the two attractors retained through modification. On average, 10.3% of students selected one of the removed options (ranging from 1 to 22%). This suggests that although attractor response rates changed from item to item, overall there was no change in the selection of attractors.

On the reading test, the results were more varied, where on average, attractor response rates decreased by about 1.2%, but varied more significantly across items (SD = 7.3). Again, the attractor response rates changed very little (with fewer students selecting the retained attractors). Table 4 contains the summary statistics for this change on both tests.

Table 4

Summary Statistics of Change in Attractor Response Rates from Original to Modified Format

	Minimum	Maximum	Mean	SD
Mathematics	-12%	11%	-0.3%	4.9%
Reading	-35%	22%	-1.2%	7.3%

As a final analysis, I examined the difference in response rates between the two retained attractors (recall that each item has one correct option and two attractors). The question is whether the difference in attractor response rates changed given the presence (original format) or absence (modified format) of one of the options. To estimate this, the difference in attractor response rates was computed (if A and B are the two attractors, % responding to A minus % responding to B). I computed the difference between the attractor response rate differences for each format (difference in attractor response rates for modified version minus original version).

|Modified % Selecting A – % Selecting B| – |Original % Selecting A – % Selecting B|

On the mathematics test, the average change in response rate differences was 1%, suggesting that overall, the difference in response rates of the two attractors did not change when a third attractor was removed. On the reading test, the average change in response rate differences was less than 1%. However, across items, there were substantial changes in attractor response rate differences, where some differences in attractor response rates decreased by up to 37% or increased as much as 27%. Table 5 contains the summary statistics for this change in response rate differences.

Table 5

Summary Statistics of Change in Differences in Attractor Response Rates from Original to Modified Format

	Minimum	Maximum	Mean	SD
Mathematics	-20%	11%	1.0%	6.7%
Reading	-37%	27%	0.9%	9.7%

In mathematics, there was a small to moderate relation between improvement in attractor discrimination and an increase in the selection rate of attractors (r = .39). This suggests one of two things: (1) that the item attractor discrimination improved as result of attracting more of the right students (lower ability students); or (2) that because attractor response rates increased, this provided more variation in responses that resulted in a higher correlation with overall performance making the attractors more discriminating.

However, this relation was not present in reading (r = .06), even though the variation in differences in attractor response rates was much greater, providing for more variance to estimate the correlation. This suggests that in mathematics, the relation was a function of the attractors attracting more of the right students. This draws strong attention to the need to make attractors plausible and attractive to students with lower ability – those students with misconceptions and reasoning errors. This of course requires the attractors to be based on misconceptions and reasoning errors.

Discussion

The present study has provided useful information regarding the role of the item response attractor because of the experimental design of the forms and randomized test administration. The experimental modifications were designed to make the items more accessible to students

with disabilities, and in doing so, improve the measurement qualities of the items. This particular package of modifications was expected to make the items somewhat easier, more discriminating, and result in more effective distractors such that they would be selected at a higher rate and more effectively discriminate between students with higher and lower ability.

Modified items were easier overall, with an average increase in correct response of 6% for mathematics and 10% for reading. Item discrimination did not fare as well, where modified items resulted in a slight decrease in discrimination of about .05 for mathematics and .01 for reading – mostly negligible. Overall, 11 of the 39 items increased in discrimination for mathematics and 17 items increased in discrimination for reading.

The key analyses for this project were at the attractor level, assessing the effect of modification on the functioning of the attractors. For both mathematics and reading, the retained attractors became more discriminating. In mathematics, 79% of all attractors across the 39 items became more discriminating (an average of .04) and 71% of all attractors in reading became more discriminating (an average of .06). The response rates to attractors did not change in either mathematics (lower by 0.3%) or reading (lower by 1.2%). Finally, differences in response rates to the two attractors within an item did not change either, where the change was 1% or less for both mathematics and reading.

The interesting analyses in this study include the examination of attractor functioning. Attractor functioning is difficult to assess as most of the statistics typically reported are affected by the difficulty of the item. For example, if the item is very easy, then the response rates to the attractors is naturally low. If there are few respondents selecting a particular attractor, then the correlation between the selection of the attractor and the total score will be affected by the limited variation due to low selection rates. Nonetheless, the trends in these results are of interest.

Closer examination of item characteristics is important in evaluating the overall effect of modification on item discrimination as results varied a great deal across items. Similarly, several items actually became more difficult following modification; these items should be reviewed as well.

There are limitations in this study that are a function of the overall quality of some items. The items employed in this study were selected based on prior item statistics. As we all know, item functioning can be a local characteristic. There were several items that performed poorly with this particular sample. In mathematics, there were 13 items with discrimination values less than .30 and 3 with values less than .20; in modified versions these counts were 20 and 6 respectively. In reading, there were 14 items with discrimination values less than .30 and 6 items with values less than .20.

One reading item in particular was troubling, where in the original format had an 18% correct response rate with discrimination of .06; in modified version this item had a correct response rate of .19 and a discrimination value of -.09, certainly not good for any purpose. This particular item was a vocabulary item employing a word that has several meanings depending on the syllable of emphasis. Although it was used in a sentence, the item required students to identify one of the alternate meanings if the stress was on the opposite syllable. Modification is not always an answer to poorly conceived items, where good item writing trumps any potential set of modifications.

A reorientation in terms of item development will help. Asking item writers or editors to attend to the attractors requires explicit attention to the attraction aspect of the incorrect options. Asking ourselves: “Is this an effective attractor?” rather than “distractor” will improve our ability to be explicit about the intent of the option and characteristics of the students to whom it attracts.

Improvements in item writing and item modifications that strive to make items and tests more widely accessible will contribute even more to the effectiveness of our attractors.

References

Ascalon, M.E., Meyers, L.S., Davis, B.W., & Smits, N. (2007). Distractor similarity and item-stem structure: Effects on item difficulty. *Applied Measurement in Education, 20*(2), 153-170.

Beddow, P.A., Kettler, R.J., & Elliott, S.N. (2008). *Test Accessibility and Modification Inventory*. Unpublished instrument. Peabody College of Vanderbilt University.

Center for Universal Design. (1997). *The principles of universal design*. Retrieved August 4, 2008 from http://www.design.ncsu.edu/cud

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load.* San Francisco, CA: Pfeiffer.

Ebel, R.L. (1951). Writing the test item. In E. F. Linquist (Ed.), *Educational measurement* (pp. 185–249). Washington, DC: American Council on Education.

Elliott, Kettler, Beddow, Kurz, Compton, McGrath, Bruen, Hinton, Palmer, Rodriguez, Bolt, Roach (under review). Using modified items to test students with and without persistent academic difficulties: Effects on groups and individual students. *Exceptional Children*.

Haladyna, T.M. (1999). *Developing and validating multiple-choice test items* (2nd ed.). Mahwah, NJ: Erlbaum.

Haladyna, T.M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Erlbaum

Haladyna, T.M., & Downing, S.M. (1988, April). *Functional distractors: Implications for test-item writing and test design*. Paper presented at the annual meeting of the AERA, New Orleans, LA.

Haladyna, T.M., & Downing, S.M. (1989a). A taxonomy of multiple-choice item writing rules. *Applied Measurement in Education, 2*, 37–50.

Haladyna, T.M., & Downing, S.M. (1989b). Validity of a taxonomy of multiple-choice item writing rules. *Applied Measurement in Education, 2*, 51–78.

Haladyna, T.M., & Downing, S.M. (1993). How many options is enough for a multiple choice item? *Educational and Psychological Measurement, 53*, 999–1010.

Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*, 309–334.

Kettler, R.J., Rodriguez, M.R., Bolt, D.M., Elliott, S.N., Beddow, P.A., & Kurz, A. (under review). Modified multiple-choice items for alternate assessments: Reliability, difficulty, and the interaction paradigm. *Applied Measurement in Education.*

McMorris, R. F., Boothroyd, R. A.,&Pietrangelo, D. J. (1997). Humor in educational testing: A review and discussion. *Applied Measurement in Education, 10,* 269–297.

Rodriguez, M.C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice, 24*(2), 3-13.

Appendix

Mathematics Item Response Frequencies

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
2	A		10	4	-.17			
	B	*	161	69	.38	323	68	.4
	C		36	16	-.23	99	21	-.25
	D		25	11	-.18	53	11	-.27
3	A		27	12	-.28	63	13	-.28
	B	*	145	62	.48	315	66	.38
	C		48	21	-.26	97	20	-.21
	D		13	6	-.17			
4	A	*	139	60	.37	320	67	.35
	B		24	10	-.25			
	C		39	17	-.21	88	19	-.34
	D		31	13	-.08	67	14	-.1
5	A		35	15	-.1	95	20	-.25
	B		31	13	-.23	104	22	-.26
	C	*	126	54	.43	276	58	.41
	D		40	17	-.27			
6	A		95	41	-.26	158	33	-.27
	B		26	11	-.2	108	23	-.33
	C	*	94	40	.49	209	44	.53
	D		18	8	-.2			
7	A		53	23	-.24	122	26	-.31
	B	*	123	53	.48	283	60	.43
	C		41	18	-.25	68	14	-.21
	D		16	7	-.17			
8	A		59	25	-.36	90	19	-.33
	B	*	102	44	.58	257	54	.49
	C		43	18	-.19	128	27	-.26
	D		29	12	-.19			

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
9	A		19	8	-.22	76	16	-.14
	B		84	36	-.3	174	37	-.4
	C	*	111	48	.55	224	47	.49
	D		19	8	-.25			
10	A		50	21	-.18	129	27	-.32
	B		29	12	-.3			
	C	*	130	56	.53	299	63	.46
	D		24	10	-.3	47	10	-.27
11	A		19	8	-.13			
	B	*	121	52	.27	260	55	.32
	C		53	23	-.05	148	31	-.22
	D		38	16	-.21	66	14	-.17
12	A		80	34	-.32	209	44	-.42
	B		30	13	-.24	56	12	-.2
	C		11	5	-.16			
	D	*	112	48	.54	210	44	.55
13	A		29	12	-.29	60	13	-.27
	B		128	55	-.14	275	58	-.16
	C	*	63	27	.47	140	29	.37
	D		13	6	-.19			
14	A		49	21	-.31			
	B		60	26	-.19	121	26	-.38
	C	*	94	40	.46	266	56	.42
	D		30	13	-.06	87	18	-.11
15	A		24	10	-.26			
	B		14	6	-.25	49	11	-.38
	C		48	20	-.39	82	18	-.36
	D	*	156	64	.61	333	72	.56
16	A		9	4	-.16			
	B	*	177	73	.52	357	77	.54
	C		47	19	-.4	93	20	-.46
	D		10	4	-.23	13	3	-.24

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
17	A	*	86	35	.4	246	53	.47
	B		51	21	-.28	107	23	-.41
	C		53	22	-.19	112	24	-.14
	D		53	22	.01			
18	A	*	159	65	.57	334	72	.48
	B		41	17	-.41	91	20	-.38
	C		33	14	-.29	40	9	-.23
	D		10	4	-.09			
19	A		48	20	-.21	105	23	-.19
	B		35	14	-.2	103	22	-.25
	C	*	107	44	.43	257	55	.37
	D		52	21	-.14			
20	A	*	121	50	.61	282	61	.54
	B		54	22	-.22	98	21	-.32
	C		39	16	-.34	85	18	-.35
	D		29	12	-.26			
21	A		22	9	-.25			
	B		25	10	-.24	67	14	-.25
	C	*	120	49	.32	195	42	.32
	D		76	31	-.03	202	44	-.15
22	A		39	16	-.26			
	B		59	24	-.23	127	27	-.32
	C		41	17	-.25	93	20	-.29
	D	*	104	43	.58	244	53	.52
23	A		18	7	-.24			
	B		35	14	-.39	66	14	-.38
	C	*	146	60	.56	284	61	.49
	D		44	18	-.18	114	25	-.24
24	A		70	29	-.23	130	28	-.3
	B		63	26	-.09	133	29	-.18
	C	*	95	39	.44	201	43	.43
	D		15	6	-.28			

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
25	A		19	8	-.23	67	14	-.26
	B	*	138	57	.42	302	65	.41
	C		66	27	-.18	96	21	-.26
	D		20	8	-.25			
26	A	*	160	66	.58	372	80	.31
	B		42	17	-.37	51	11	-.28
	C		28	12	-.26	41	9	-.13
	D		13	5	-.23			
27	A		55	23	-.18	122	26	-.23
	B		48	20	-.32	70	15	-.31
	C	*	121	50	.48	272	59	.44
	D		18	7	-.14			
28	A		55	24	-.21	140	29	-.26
	B		43	19	-.23	168	35	-.19
	C	*	84	36	.52	167	35	.44
	D		50	22	-.16			
29	A		14	6	-.26			
	B		63	27	-.32	142	30	-.47
	C		46	20	-.08	89	19	-.24
	D	*	109	47	.47	244	51	.62
30	A	*	120	52	.5	231	49	.56
	B		60	26	-.21	159	33	-.34
	C		49	21	-.34	85	18	-.32
	D		3	1	-.18			
31	A		39	17	-.22	97	20	-.23
	B	*	129	56	.22	287	60	.26
	C		51	22	.02	92	19	-.08
	D		12	5	-.16			
32	A		61	26	.02	148	31	-.05
	B	*	67	29	.37	169	36	.34
	C		82	35	-.21	159	33	-.3
	D		22	9	-.24			

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
33	A	*	97	42	.39	203	43	.5
	B		52	22	-.18	157	33	-.26
	C		43	19	-.2	116	24	-.29
	D		40	17	-.11			
34	A		34	15	-.15	90	19	-.3
	B	*	121	52	.58	317	67	.46
	C		39	17	-.32	65	14	-.29
	D		38	16	-.32			
35	A		38	16	-.32	76	16	-.31
	B		97	42	-.06	241	51	-.11
	C	*	77	33	.4	159	33	.35
	D		19	8	-.14			
36	A	*	163	70	.49	350	74	.44
	B		39	17	-.33	96	20	-.33
	C		19	8	-.22	30	6	-.27
	D		11	5	-.2			
37	A		36	16	-.17	79	17	-.22
	B		48	21	-.2	142	30	-.4
	C	*	112	48	.54	254	53	.54
	D		36	16	-.34			
38	A	*	163	71	.49	376	79	.41
	B		47	20	-.35	78	16	-.35
	C		16	7	-.23	22	5	-.18
	D		5	2	-.17			
39	A		50	22	.09			
	B	*	96	41	.41	238	50	.37
	C		34	15	-.25	131	28	-.33
	D		52	22	-.35	106	22	-.08
40	A		35	15	-.28	102	21	-.33
	B	*	111	48	.52	266	56	.47
	C		51	22	-.21	108	23	-.23
	D		35	15	-.19			

Reading Item Response Frequencies

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
2	A	*	182	76	.57	396	86	.49
	B		24	10	-.35	30	7	-.36
	C		13	5	-.27			
	D		20	8	-.27	33	7	-.31
3	A		24	10	-.28	46	10	-.32
	B	*	166	69	.56	345	75	.48
	C		40	17	-.33	68	15	-.3
	D		10	4	-.27			
4	A		58	24	-.27	46	10	-.46
	B		24	10	-.37			
	C		21	9	-.23	35	8	-.36
	D	*	136	57	.59	377	82	.61
5	A		26	11	-.17	79	17	-.36
	B		77	32	-.11			
	C	*	78	33	.3	310	68	.59
	D		59	25	-.08	68	15	-.40
6	A		8	3	-.19	22	5	-.35
	B		27	11	-.33	40	9	-.37
	C		10	4	-.29			
	D	*	194	81	.5	395	86	.53
7	A	*	186	78	.6	377	82	.51
	B		30	13	-.36	61	13	-.43
	C		14	6	-.32	21	5	-.24
	D		9	4	-.29			
8	A		40	17	-.06	68	15	-.1
	B		45	19	-.21	102	22	-.44
	C	*	132	55	.36	289	63	.45
	D		23	10	-.24			

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
9	A		47	20	-.2	170	37	-.31
	B		70	29	-.17			
	C		41	17	-.25	71	15	-.16
	D	*	82	34	.52	218	47	.42
10	A		76	32	.01	164	36	.05
	B		52	22	-.18	107	23	-.32
	C		36	15	-.16			
	D	*	76	32	.28	188	41	.23
11	A		70	29	-.17	135	29	-.12
	B		44	18	-.21			
	C	*	85	35	.43	214	47	.3
	D		41	17	-.13	110	24	-.23
12	A		63	26	-.13	84	18	-.2
	B	*	73	30	.41	229	50	.38
	C		82	34	-.18	145	32	-.23
	D		22	9	-.16			
13	A		47	20	-.33	104	23	-.4
	B		17	7	-.23	31	7	-.26
	C		8	3	-.19			
	D	*	168	70	.49	324	71	.51
14	A		20	8	-.25			
	B		34	14	-.31	54	12	-.38
	C		45	19	-.27	93	20	-.25
	D	*	141	59	.57	312	68	.47
15	A	*	179	80	.54	391	83	.47
	B		11	5	-.31	20	4	-.34
	C		30	13	-.38	62	13	-.33
	D		4	2	-.18			
16	A		27	12	-.26			
	B		39	17	-.25	82	17	-.26
	C		20	9	-.25	39	8	-.33
	D	*	137	61	.51	350	74	.43

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
17	A		57	26	-.17			
	B		32	14	-.27	130	27	-.45
	C	*	95	43	.49	240	51	.55
	D		39	17	-.19	103	22	-.19
18	A	*	105	47	.45	287	61	.41
	B		42	19	-.19	111	23	-.36
	C		57	26	-.17	75	16	-.13
	D		19	9	-.27			
19	A		66	30	-.22			
	B	*	111	50	.49	349	74	.61
	C		22	10	-.25	67	14	-.42
	D		24	11	-.24	57	12	-.38
20	A		31	14	.05	106	22	-.24
	B		28	13	-.24	66	14	-.26
	C		35	16	-.43			
	D	*	130	58	.44	301	64	.4
21	A	*	143	64	.48	357	76	.51
	B		32	14	-.23	74	16	-.33
	C		24	11	-.29	41	9	-.35
	D		25	11	-.19			
22	A		131	58	-.31	123	26	-.38
	B		30	13	-.1	85	18	-.23
	C	*	53	24	.48	265	56	.51
	D		10	4	-.07			
23	A		42	19	-.26			
	B	*	139	62	.38	293	62	.42
	C		23	10	-.2	75	16	-.31
	D		20	9	-.09	105	22	-.23
24	A		16	7	-.19	49	10	-.35
	B		16	7	-.29	52	11	-.37
	C	*	140	63	.44	371	79	.54
	D		52	23	-.22			

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
25	A		23	10	-.2	61	13	-.3
	B	*	137	61	.53	233	49	.46
	C		46	21	-.3	179	38	-.26
	D		18	8	-.29			
26	A		6	3	-.1	30	6	-.31
	B	*	207	92	.23	427	90	.4
	C		3	1	-.22	16	3	-.24
	D		8	4	-.12			
27	A		85	38	.23	111	24	-.03
	B		69	31	-.26	272	58	-.03
	C	*	40	18	.21	89	19	.07
	D		30	13	-.21			
28	A	*	162	69	.57	350	75	.55
	B		23	10	-.25	61	13	-.4
	C		31	13	-.39	54	12	-.32
	D		18	8	-.21			
29	A		31	13	-.29			
	B	*	173	75	.53	383	82	.54
	C		12	5	-.26	33	7	-.31
	D		16	7	-.29	50	11	-.41
30	A	*	128	55	.31	347	74	.41
	B		55	24	-.22	70	15	-.29
	C		28	12	-.12	49	11	-.25
	D		23	10	-.08			
31	A	*	133	57	.52	332	71	.5
	B		42	18	-.17			
	C		24	10	-.29	65	14	-.4
	D		34	15	-.3	68	15	-.24
32	A		22	9	-.34			
	B		39	17	-.36	82	18	-.45
	C	*	157	67	.61	341	74	.6
	D		15	6	-.2	40	9	-.33

			Original Format			Modified Format		
Item	Option	Key	Count	%	PTBS	count	%	PTBS
33	A		23	10	-.33	54	12	-.41
	B	*	62	27	.38	224	48	.33
	C		48	21	-.18	186	40	-.07
	D		100	43	.01			
34	A		9	4	-.26	11	2	-.19
	B		7	3	-.19			
	C		12	5	-.13	33	7	-.32
	D	*	206	88	.35	418	90	.38
35	A		38	16	-.28	50	11	-.38
	B	*	138	59	.57	390	84	.52
	C		27	12	-.32	25	5	-.33
	D		31	13	-.22			
36	A		37	16	-.24	77	17	-.2
	B		80	34	-.29	143	31	-.32
	C		8	3	-.24			
	D	*	109	47	.54	245	53	.44
37	A	*	112	48	.47	237	51	.58
	B		82	35	-.27	175	38	-.42
	C		26	11	-.24	54	12	-.28
	D		13	6	-.13			
38	A		17	7	-.37	41	9	-.33
	B		38	16	-.37	78	17	-.43
	C	*	170	73	.61	347	74	.58
	D		9	4	-.22			
39	A	*	135	58	.49	272	59	.54
	B		24	10	-.21	68	15	-.28
	C		43	18	-.25	123	27	-.38
	D		31	13	-.24			
40	A	*	121	52	.49	258	55	.37
	B		36	15	-.22	91	20	-.29
	C		38	16	-.22	116	25	-.15
	D		39	17	-.23			