

**A Comparison of SSI and Non-SSI Performance  
On State NAEP Mathematics Assessment Items**

Norman Webb, Allan Cohen, and Taehoon Kang  
University of Wisconsin  
Madison, WI

April 2, 2002

Paper Presented at the Annual Conference of the American Educational Research Association,  
New Orleans, LA, April, 2002.

Running Head: DIF Analysis of SSI

The research presented in this paper was supported by the National Science Foundation under Grant No. REC-98874171 and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison. Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the views of the supporting agencies.

## Abstract

The impact of curricular reform efforts that are incorporated into the Statewide Systemic Initiatives (SSIs) program were examined at the item level using data from the 1990, 1992, and 1996 NAEP state administrations. The item-level data from the State NAEP for grade 8 from 1990, 1992, and 1996 and for grade 4 from 1992, and 1996 were examined applying methodology developed for the study of differential item functioning (DIF) under Item Response Theory. The results suggest that there are some important differences in item functioning between SSI and non-SSI states. Eighth grade Students in SSI states in 1992 and 1996 appear more likely to perform better than those in non-SSI states on items reflecting curricular reforms associated with the NCTM *Standards*. At grade 4, the earlier differences between SSI states and non-SSI states declined, indicating that SSI states became more comparable with the non-SSI states in the underlying constructs being tested.

Index Terms: Item Response Theory, Differential Item Functioning, National Assessment of Educational Progress, Statewide Systemic Initiative.

## **A Comparison of SSI and Non-SSI Performance on State NAEP Mathematics Assessment Items**

The impact of the curricular reforms that are part of the Statewide Systemic Initiatives (SSIs) funded by the National Science Foundation (NSF) was examined at the item level. The focus of this inquiry was to determine whether curricular effects might be present in these data that could be ascribed to participation in the SSI program. In order to investigate this question, methods used for detection of differential item functioning (DIF) typically associated with item bias were used in this study. The assumption is that DIF is a nuisance dimension intruding on the ability intended to be measured (Ackerman, 1992; Roussos & Stout, 1996). In the context of this study, the focus was on whether the construct being measured in the SSI and non-SSI samples was the same. The presence of DIF items represents a difference in the construct between what is being measured and the items identified.

An underlying vision of NSF for the SSIs was that all students could achieve the ambitious mathematics outcomes described by the National Council of Teachers of Mathematics *Curriculum and Evaluation Standards for School Mathematics* (1989) (Zucker, Shields, Adelman, Corcoran, & Goertz, 1998). The NCTM *Standards* emphasize the importance of reasoning, communication, and problem solving for the in-depth learning of mathematics, along with learning to compute with numbers, analyze data, use geometry, and apply the principles of algebra. Mathematics reform strategies in many SSI states, and non-SSI states, included emphasis on these factors. But because of the strong emphasis by NSF on this content, it is reasonable to investigate whether SSI states showed greater improvement on content that was given increased attention in the NCTM *Standards*. It was hypothesized that NAEP items reflecting outcomes in such areas might function differently in SSI and non-SSI states. In particular, examinees in states emphasizing reform would be expected to perform qualitatively differently on such items than examinees in states that did not.

Item Response Theory (IRT) (Hambleton & Swaminathan, 1985; Lord, 1980) models for NAEP mathematics items (Allen, Jenkins, Kulick, & Zelenak, 1997) were used to analyze items administered as part of the State NAEP assessment. Likelihood ratio tests for differential item functioning (DIF) were used to compare the performance of students in SSI and non-SSI states. The focus of this study was to determine whether patterns existed in these data that might indicate the influence of SSI participation.

### **Methods**

#### **Data**

Data for this study were taken from the State NAEP Assessment results for grade 8 for 1990, 1992, and 1996 and for grade 4 for 1992 and 1996. The sample sizes and ethnic group composition of these samples are given in Table 1. DIF comparisons between SSI and non-SSI samples were conducted on these data. It is evident that the sample sizes are quite large. The states included in the analyses are those states for which data were available for all three years for grade 8 and for both years for grade 4. There were 17 states in the SSI sample and 11 states in

the non-SSI sample (see Table 2). Included in these 17 SSI states are four states that received funding from NSF, but had the funding withdrawn after one or two years, well before the end of the five-year grant period. Although NSF deemed that these four states had not demonstrated adequate progress towards systemic reform to receive continued funding, there is a possibility that even being a part of the SSI program for a short time may have resulted in some improvement in student achievement that could be related in part to NSF's SSI program. Therefore, we retained these states in the DIF analysis.

-----  
Insert Tables 1 and 2 About Here  
-----

The items in the State NAEP assessment are administered in blocks. Each student receives only a portion of the total block of items available for that year and grade level. In most cases, students received three blocks of between 10 and 15 items each. Some blocks consisted of multiple-choice items, some consisted of constructed response items and others contained both multiple-choice and constructed-response items. Some of the constructed-response items were considered short enough to fit with a 2-parameter IRT model (2PL) while other constructed response items were fit with a graded-response model. The IRT models used in this study were those described for each item in Allen et al. (1997). The numbers of items and blocks by grade and year are given in Table 3.

-----  
Insert Table 3 About Here  
-----

It can be seen in Table 3 that the number of blocks is consistent across years except for grade 8, 1990, for which ten blocks were administered but only seven blocks for which data were available. Thirteen blocks were provided in each of the other grade-by-year administrations. The number of items changed, however, somewhat markedly. The largest percentage of items for any one year was allocated to measuring content area 1 (Number and Operations), although, from 1990 to 1996, the proportion of grade 8 items in these seven blocks which were allocated to measuring Number and Operations declined from 33% in 1990 to 28% in 1996. Over this period, the proportion of grade 8 items in these same blocks allocated to measuring Algebra and Functions increased slightly, from 18% in 1990 to 20%. The proportion of items allocated to the other three content areas in these seven blocks remained fairly constant from 1990 to 1996. Over time, process areas 1 (conceptual understanding) and 3 (problem solving) received more emphasis than procedural knowledge and in 1996 had the highest proportion of items across both grades. The most frequent type of item was clearly the multiple-choice item.

### **NAEP Item Categories**

The item categories for the State NAEP assessment are given in Table 4 by grade and by year. NAEP categorizes items as belonging to one of five content areas for 4<sup>th</sup> and 8<sup>th</sup> grade mathematics: (1) Number and Operations, (2) Measurement, (3) Geometry, (4) Data Analysis,

Statistics, and Probability, and (5) Algebra and Functions. Four process categories are identified for these same items: (1) conceptual understanding, (2) procedural knowledge, (3) problem solving (these are multiple-choice items), and (4) problem solving with extended response (typically, these ask students to show their work). The content and process breakdowns by item format are given in Tables 4 and 5 for grades 4 and 8, respectively, for each year. When possible, we have used this designation to identify patterns of DIF in the SSI versus non-SSI comparisons.

-----  
Insert Tables 4 and 5 About Here  
-----

### **Detection of DIF**

Detection of DIF was done in the context of IRT (Hambleton & Swaminathan, 1985; Lord, 1980) using the likelihood ratio test for DIF (Thissen, Steinberg & Wainer, 1988, 1993). The IRT models used in this study were the 2- and 3-parameter logistic models and the graded response models as described by Allen et al. (1997) for each of the items. Previous research has shown that the likelihood ratio test for DIF controls Type I error at the item level for the 2- and 3-parameter IRT dichotomous models (Cohen, Kim & Wollack, 1996) and for the graded response models (Kim & Cohen, 1998). The likelihood ratio test for DIF was conducted using output from the computer program MULTILOG (Thissen, 1991). All DIF results were obtained within blocks for both dichotomous and constructed-response items.

### **Results**

The presence of DIF items in this study indicates a difference in the underlying construct being measured in the SSI and non-SSI groups. Our focus is on the use of DIF items as an indication of the impact of the SSI initiative on the mathematics construct in grades 4 and 8. In this section, we discuss some of the possible patterns that are present in the DIF results. Grade 4 DIF items are identified in Table 6 and grade 8 items in Table 7 for each of the years for which data were analyzed.

-----  
Insert Tables 6 and 7 About Here  
-----

The clear majority of DIF items in Tables 6 and 7 are multiple-choice (Type 2). It is interesting to note that, although there were relatively few extended open-ended items (Type 3) in any of the grade-by-year combinations, none were identified as DIF for grade 4 and only one in grade 8 (Item 9, Block M14, 1996). SSI status (i.e., whether the state was included in the SSI or non-SSI samples) was not found to be related to content classification in Tables 6 or 7. That is, the content classification provided by NAEP was not related to SSI status. Further, process, as identified by NAEP, was related to SSI status in only one sample, grade 8 in 1990 (see Table 7a). It appears from Table 7a that DIF items in process category 2 (procedural knowledge) and 3 (problem solving) were more often in the non-SSI sample. Eleven of the 13 DIF items in process category 2 or 3 favored the non-SSI group.

The pattern of DIF between SSI and non-SSI was different for grades 4 and 8. It is clear from Tables 6a and 6b that the number of items identified in grade 4 as functioning differentially between the SSI and non-SSI states is much smaller in 1996 than in 1992. The patterns for each grade by year combination are shown graphically in Figure 1.

-----  
Insert Figure 1 About Here  
-----

### **Grade 4 Items**

More DIF items were detected for grade 4 in 1992 than in 1996 (see Figure 1): 26 of the 156 items (17%) were DIF in 1992, while only 9 of 144 items (6%) were DIF in 1996. In 1992, grade 4 students in SSI states showed better performance in seven Number and Operations items compared to students with the same mathematical ability in non-SSI states. Students in non-SSI states showed better performance in 1992 on only one Number and Operations item. Grade 4 students in non-SSI states performed higher on seven Measurement items compared to students with the same ability in SSI states. Students in SSI states performed better in 1992 on only two Measurement items.

Eight DIF items measuring conceptual understanding were observed in the grade 4, 1992, assessments to favor the SSI sample (Table 8). Four items measuring procedural knowledge and three items measuring problem solving also were found to favor the SSI sample. DIF items favored grade 4 students in the non-SSI sample on five conceptual understanding items, one procedural knowledge item, and five problem-solving items compared to the students with same ability in the SSI sample (Table 9).

-----  
Insert Tables 8 and 9 About Here  
-----

In 1992, grade 4 students in SSI states participating in the State NAEP, when compared to students of equal ability in non-SSI states participating in the state NAEP, performed better on a higher number of Number and Operations items at all three process levels. Students in non-SSI states performed better on Measurement items, particularly those categorized as conceptual understanding and problem solving. The number of DIF items categorized in the other three content areas was too small to indicate any distinguishing characteristics between the two groups.

In addition to considering the number of items in different categories, a sample of items was administered in both testing years. This provided an opportunity to investigate whether the same item changed in DIF status from 1992 to 1996. Only nine DIF items were found in 1996. This is a much lower number than the 26 DIF items detected in 1992. It appears that some portion of the difference in the constructs observed in grade 4 in 1992 disappeared in 1996.

The State NAEP had 13 different blocks of mathematics items. Eight blocks administered in 1992 (blocks 3, 4, 6, 8, 9, 10, 11, and 13) were administered again in 1996. Of the 97 items

only in these eight blocks, 16 DIF items were found in 1992 and eight DIF items in 1996 (Tables 6a and 6b). The pattern of DIF for these two administrations is given in Table 6c. In both years, the number of DIF items of those administered each year was evenly distributed between SSI and non-SSI states. In 1992, eight of the 16 DIF items favored the SSI sample. In 1996, four of the nine DIF items favored the SSI states. Of the eight DIF items that favored students from the SSI states in 1992, none were detected as functioning differentially in 1996. What initially differentiated the constructs measured in the two groups on these items disappeared.

Also, of eight items favoring the non-SSI sample in 1992, six were non-DIF in 1996. One of the two DIF items favoring the non-SSI students in 1992 (item 7 of block M6) favored non-SSI students in 1996 and one (item 6 of block M10) favored the SSI students in 1996. Of the 81 non-DIF items in 1992 from those blocks that were repeated in 1996, two items were detected with DIF favoring the SSI sample and four items were detected with DIF favoring the non-SSI sample. The only item that switched between the two groups was a non-SSI-favored item in 1992 that became a SSI-favored item in 1996. This item (item 6 of block M10) was an open-ended item assessing Measurement (comparison of areas of two shapes).

The pattern of DIF items by topic and process does suggest some differences at grade 4 in 1992 between the performance of students in SSI states and non-SSI states. However, in both number of DIF items and their categorization by topic and process, there is little indication of any such differences at grade 4 in 1996. Four more DIF items favored the SSI states than favored the non-SSI states in 1992. This number is not large. What is more noticeable, however, is the distribution of DIF items favoring SSI states included a large proportion of Number and Operations items (solve a problem with numbers and place value) with only two Measurement items (area and interpret speedometer). The DIF items in 1992 favoring the non-SSI states included a large proportion of Measurement items (length, area, and volume) and only one Number and Operations item (use multiplication).

In 1992, the overall composite average score by SSI states was five points lower than by non-SSI states. The proportion of White students was about 10% lower in SSI states than non-SSI states included in the analysis (Webb, Kane, Kaufman, and Yang, 2001). The expectation was that there would be relatively few items in 1992 that would favor the SSI states over the non-SSI states. The number of DIF items that did favor the SSI states, however, was surprising. It was further interesting that these were related to students' solving number problems. The differences between the two groups of states diminished from 1992 to 1996 and were not as apparent in 1996. This would suggest that the SSI and non-SSI states became more similar, not only in reducing the performance gap in the average scale score by one point, but also in the type of items that distinguished the performance between the two groups.

## **Grade 8 Items**

The number of DIF items for grade 8 was largest in 1990. Fifty of 137 items (36%) were identified as DIF in 1990. In 1992, 25 of 183 items (14%) were identified as DIF. In 1996, 21 of 162 items (13%) were detected as DIF. From a study of randomly distributed items, it was found that about 5% of the items could have been DIF by chance. Across all three of the years analyzed

in this study, the number of DIF items clearly exceeded what could be expected by chance. In conjunction with the decline in the number of DIF items over this period, it appears that the differences in the underlying construct of mathematics achievement measured by the State NAEP items were decreasing during this period.

In 1990, 11 items measuring Number and Operations favored the SSI states when compared to students with same mathematical ability in non-SSI states (Tables 7a and 10). There were also 11 Number and Operations items that favored the non-SSI sample in 1990 (Tables 7a and 11). In 1992, 11 items were found favoring the SSI sample (Tables 7b and 12) and 14 favoring the non-SSI sample (Tables 7b and 13). In 1996, 15 items favored the SSI states and six items favored the non-SSI states (Tables 7c, 14 and 15). The number of DIF items decreased in the non-SSI states from 1990 to 1996. The number of DIF items favoring the SSI states that measured Number and Operations and Algebra and Functions decreased from 1990 to 1992, but then increased in 1996 in the SSI states. It does appear that there was a slight SSI impact on Number and Operations and on Algebra and Functions by 1996.

-----  
Insert Tables 10 and 11 About Here  
-----

For grade 8 students in 1990, 15 DIF items measuring conceptual understanding were found favoring the SSI sample (Tables 7a and 10). An additional 7 items measuring procedural knowledge and five items measuring problem solving were also found favoring SSI students when comparing grade 8 students with students of the same ability as students in non-SSI states. Also in 1990, nine conceptual understanding items, ten procedural knowledge items, and four problem-solving items favored non-SSI students (Tables 7a and 11). The numbers of DIF items favoring either SSI or non-SSI samples decreased from 1990 to 1996. The number of SSI favored items measuring conceptual understanding decreased from 1990 to 1992 but then increased from 1992 to 1996. In contrast, the number of conceptual understanding DIF items favoring non-SSI states decreased steadily over the three testing periods. There does appear to be some positive SSI impact, therefore, over the three NAEP administrations in conceptual understanding for grade 8.

The State NAEP had 13 blocks in mathematics for grade 8; eight of the blocks (blocks 3, 4, 6, 8,9, 10, 11, and 13) administered in 1996 had been administered in 1992. The pattern of DIF and non-DIF items for these two administrations is given in Table 7d. Of the 114 items in these eight blocks, 18 were DIF items in 1992 and 18 were DIF items in 1996. Seven of the nine DIF items favoring the SSI sample in 1992 were non-DIF items in 1996. Six of the nine 9 DIF items favoring the non-SSI sample in 1992 also became non-DIF items in 1996. Of the 96 non-DIF items in 1992, 10 were DIF items favoring the SSI sample and three were DIF items favoring the non-SSI sample in 1996. Ten non-DIF items in 1992 became DIF items favoring SSI states in 1996, two non-SSI favored DIF items in 1992 changed into SSI-favored DIF items, and one SSI favored DIF item in 1992 remained a SSI-favored DIF item in 1996. Thus, 13 of the 114 repeated items, 11%, were SSI-favored DIF items in 1996. This is compared to five of the 114 repeated items, 4%, that were DIF items favoring the non-SSI in 1996.



For grade 8, over the three testing times—1990, 1992, and 1996—as was observed in grade 4, the six-point gap in average student performance between the 11 non-SSI states and the 17 SSI states decreased by one scale point. The SSI states had a lower average composite score that increased by 1.2 scale points over the six years (Webb, Kane, Kaufman, and Yang, 2001). This supports the contention that the gain in performance for SSI states was related to some degree by the SSI intervention. The DIF analysis for grade 8 revealed some of the nature of this change and produced additional evidence of the positive impact on the type of items students in the SSI states could do when compared to students of comparable abilities in the non-SSI states. The decrease in the total number of DIF items over time indicates that the construct underlying student performance in the two groups of states became more similar over time, 50 DIF items in 1990 compared to 21 DIF items in 1996. The decrease in 29 items is greater than expected by chance. The number of DIF items favoring the 17 SSI states declined between 1990 and 1992, then increased in 1996. This V-shape curve differs from the steady decline in the number of DIF items favoring the non-SSI states. In 1990, the SSI states and non-SSI states varied little by topic, but the SSI states did better on six DIF items categorized as conceptual understanding than did the non-SSI states. In 1992, the SSI states and non-SSI states appeared to become more similar in terms of the DIF items favoring each group. The number of DIF items favoring the

-----  
Insert Tables 12 and 13 About Here  
-----

non-SSI states in the topics Number and Operations, Measurement, and Geometry did exceed the number of DIF items in those three topics favoring the SSI states. This may be a start of a trend. In 1996, after four years of the SSI program, the number of DIF items in two topics, Number and Operations and Algebra and Functions, favoring SSI states exceeded the number favoring the non-SSI states (Tables 14 and 15). Eight more DIF items in conceptual understanding and problem solving, plus one extended-response item, favored SSI states compared to non-SSI states. The trend for grade 8 over the three test administrations is consistent with the interpretation that SSI states made some improvements relative to the non-SSI states. This can be seen by the reduction in the number of DIF items in conceptual understanding and problem solving and an increase in the number of DIF items favoring SSI in Number and Operations and Algebra and Functions.

-----  
Insert Tables 14 and 15 About Here  
-----

## **Discussion**

The presence of DIF items in this study was interpreted as an indication of the differences in the construct being measured by the State NAEP over the period from 1990 to 1996. In both grades 4 and 8, we did find some differences that favored the SSI states and some that favored the non-SSI states. Although the majority of items on the State NAEP were multiple choice,

there was no pattern of DIF favoring either the SSI or non-SSI samples by item format. This suggests that item format (albeit predominantly multiple-choice) was not related to performance in the SSI or non-SSI samples. There were very few extended open-ended items on either the grade 4 or grade 8 examinations and only one of these items was found to function differentially, favoring the SSI sample.

Item content categories as identified by NAEP did appear to have some relationship to SSI/non-SSI status. At grade 4 in 1992 prior to the beginning of the SSI program, SSI students performed better than non-SSI students of comparable ability on Number and Operations items while non-SSI students performed better on Measurement items. At grade 8 in 1992, non-SSI students performed better on items from the more traditional grade 8 topics of Number and Operations, Measurement, and Geometry. Four years later in 1996, the SSI students performed better than students from non-SSI states with comparable abilities on Number and Operations and Algebra and Functions items. Item process was only related to SSI/non-SSI status in grade 8 in 1990. This relationship disappeared by 1992 and only reappeared in 1996 when we aggregated conceptual understanding and problem solving.

The number of DIF items decreased in both grades 4 and 8 over the periods studied. This can be seen clearly in terms of the total number of DIF items for 1990, 1992, and 1996. It is also apparent in terms of specific items that were administered both in 1992 and 1996. The decrease in DIF among these items provides some indication of the narrowing of the differences in the underlying constructs measured in the SSI and non-SSI groups. It is interesting to note the strength of these findings, given the diversity in SSI implementations among the states in the SSI sample.

Achievement differences are somewhat related to these differences in that the kinds of items favoring one group over another may be a cause of some of the differences. Other research has pointed to clear and continuing differences in achievement between SSI and non-SSI states. In the case of the SSI versus non-SSI comparisons in this paper, it appears that the construct differences (albeit not necessarily the achievement gap) may be decreasing.

This study demonstrates the viability of using DIF analysis for investigating the differences among SSI-related state reforms. The analyses were performed using the same IRT model as that used by NAEP and the NAEP item-categorization scheme and item descriptors were used to look for patterns in DIF items. At most, three point trends were analyzed in this study. An analysis using the state NAEP 2000 data would help determine whether the changes observed in 1996 really comprised a trend or only a unique variation.

## References

Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. Journal of Educational Measurement, 29(1), 67-91.

- Allen, N. L., Jenkins, F. Kulick, E., & Zelenak, C. A. (1997). Technical report of the NAEP 1996 state assessment program in mathematics. Washington, DC: U.S. Government Printing Office.
- Cohen, A.S., Kim, S.-H., & Wollack, J. A. (1996). An investigation of the likelihood ratio test for detection of differential item functioning. Applied Psychological Measurement, 20(1), 15-26.
- Hambleton, R.K., & Swaminathan, H. (1985). Item Response Theory: Principles and applications. Boston, MA: Kluwer-Nijhoff Publishing.
- Kim, S.-H., & Cohen, A.S. (1998). Detection of differential item functioning under the graded response model with the likelihood ratio test. Applied Psychological Measurement, 22(4), 345-355.
- Lord, F.M. (1980). Applications of Item Response Theory to practical testing problems. Hillsdale, NJ: Erlbaum.
- National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.
- Roussos, L., & Stout, W. (1996). A multidimensionality-based DIF analysis paradigm. Applied Psychological Measurement, 20, 355-371.
- Thissen, D. (1991). MULTILOG (version 6.0) [computer program]: Multiple, categorical item analysis and test scoring using item response theory. Chicago, IL: Scientific Software, Inc.
- Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of Item Response Theory in the study of group differences in trace lines. In H. Wainer & H.I. Braun (Eds.), Test validity. (pp. 147-169). Hillsdale, NJ: Erlbaum.
- Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P.W. Holland & H. Wainer (Eds.), Differential item functioning, (pp.67--113). Hillsdale, NJ: Erlbaum
- Webb, N. L., Kane, J., Kaufman, D., & Yang, J. (2001). *Study of the impact of Statewide Systemic Initiatives Program: Technical report to the National Science Foundation on the use of State NAEP data to assess the impact of the Statewide Systemic Initiatives*. Madison: University of Wisconsin, Wisconsin Center for Education Research.
- Zucker, A. A., Shields, P. M., Adelman, N. E., Corcoran, T. B., & Goertz, M. E. (1998). *Statewide Systemic Initiatives Programs: A report on the evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) program*. Menlo Park, CA: SRI International.

Table 1  
*Sample Sizes of NAEP State Assessments*

		Caucasian	African-American	Hispanic.	Asian and Pacific Islander	Native American	Unclassified	Total
1990 Grade 8	SSI	87551	20850	17025	3120	2868	207	131621
	Non-SSI	63864	6657	6492	6456	2145	174	85788
1992 Grade 4	SSI	83130	22512	19571	3102	2670	162	131147
	Non-SSI	63330	7737	7614	5925	2721	219	87546
1996 Grade 4	SSI	78724	21071	18111	2949	3219	195	124269
	Non-SSI	57765	8202	8208	5199	2436	216	82026
1992 Grade 8	SSI	84450	21300	17315	3189	1650	276	128180
	Non-SSI	63712	6918	5982	5859	1950	177	84598
1996 Grade 8	SSI	76055	19608	16444	3393	2025	241	117766
	Non-SSI	56757	7157	5843	5293	1712	201	76963

Table 2  
*SSI and Non-SSI States in the Sample*

<b>SSI States</b>	<b>Non-SSI States</b>
Arkansas	Alabama
California	Arizona
Colorado	Hawaii
Connecticut	Indiana
Delaware	Iowa
Florida	Maryland
Georgia	Minnesota
Kentucky	North Dakota
Louisiana	West Virginia
Michigan	Wisconsin
Nebraska	Wyoming
New Mexico	
New York	
North Carolina	
Rhode Island	
Texas	
Virginia	

Table 3  
*Item Blocks by Content and Process*

Year		1990		1992		1996	
Grade		8th	4th	8th	4th	8th	8th
Num. of items (Num. Of Blocks)		137 (7)	156 (13)	183 (13)	144 (13)	162 (13)	
<b>Content</b>	1 number & operation	46	63	58	59	46	
	2 measurement	21	29	32	25	27	
	3 geometry	26	27	36	25	31	
	4 data analysis, statistic & probability	19	20	28	17	25	
	5 algebra & functions	25	17	29	18	33	
<b>Process</b>	1 conceptual understanding	58	64	67	59	59	
	2 procedural knowledge	42	31	45	30	44	
	3 problem solving	37	56	65	48	51	
	4 problem solving extended pen quest	0	5	6	7	8	
<b>Item Type</b>	1 open ended	35	54	59	57	61	
	2 multiple choice	102	97	118	80	93	
	3 extended open ended	0	5	6	7	8	

Table 4  
Grade 4 Item Type Designations

Table 4a. Grade 4, 1992 (content by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Content				
1 (number & operation)	20	41	2	63
2 (measurement)	7	22	0	29
3 (geometry)	12	14	1	27
4 (data analysis, stat. & prob.)	9	10	1	20
5 (algebra & functions)	6	10	1	17
Total	54	97	5	156

Table 4b. Grade 4, 1992 (process by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Process				
1 (conceptual understanding)	17	47	0	64
2 (procedural knowledge)	14	17	0	31
3 (problem solving)	23	33	0	56
4 (prob. solving with ext.)	0	0	5	5
Total	54	97	5	156

Table 4c. Grade 4, 1996 (content by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Content				
1 (number & operation)	24	34	1	59
2 (measurement)	6	19	0	25
3 (geometry)	13	10	2	25
4 (data analysis, stat. & prob.)	6	9	2	17
5 (algebra & functions)	8	8	2	18
Total	57	80	7	144

Table 4d. Grade 4, 1996 (process by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Process				
1 (conceptual understanding)	18	41	0	59
2 (procedural knowledge)	13	17	0	30
3 (problem solving)	26	22	0	48
4 (prob. solving with ext.)	0	0	7	7
Total	57	80	7	144

Table 5  
Grade 8 Item Type Designations

Table 5a. Grade 8, 1990 (content by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Content				
1 (number & operation)	10	36	0	46
2 (measurement)	4	17	0	21
3 (geometry)	8	18	0	26
(data analysis, stat. & prob.)	6	13	0	19
5 (algebra & functions)	7	18	0	25
Total	35	102	0	137

Table 5b. Grade 8, 1990 (process by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Process				
1 (conceptual understanding)	12	46	0	58
2 (procedural knowledge)	13	29	0	42
3 (problem solving)	10	27	0	37
4 (prob. solving with ext.)	0	0	0	0
Total	35	102	0	137

Table 5c. Grade 8, 1992 (content by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Content				
1 (number & operation)	15	41	2	58
2 (measurement)	12	19	1	32
3 (geometry)	15	20	1	36
4 (data analysis, stat. & prob.)	10	17	1	28
5 (algebra & functions)	7	21	1	29
Total	59	118	6	183

Table 5d. Grade 8, 1992 (process by item type)

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Process				
1 (conceptual understanding)	15	52	0	67
2 (procedural knowledge)	14	31	0	45
3 (problem solving)	30	35	0	65
4 (prob. solving with ext.)	0	0	6	6
Total	59	118	6	183



Table 5, continued:

Table 5e. *Grade 8, 1996 (content by item type)*

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Content				
1 (number & operation)	14	30	2	46
2 (measurement)	10	15	2	27
3 (geometry)	15	15	1	31
4 (data analysis, stat. & prob.)	10	12	3	25
5 (algebra & functions)	12	21	0	33
Total	61	93	8	162

Table 5f. *Grade 8, 1996 (process by item type)*

Item type	1 (Open ended)	2 (Multiple choice)	3 (Ext. open ended)	Total
Process				
1 (conceptual understanding)	18	41	0	59
2 (procedural knowledge)	11	33	0	44
3 (problem solving)	32	19	0	51
4 (prob. solving with ext.)	0	0	8	8
Total	61	93	8	162

Table 6  
DIF Items by Content, Process, and Item Type

Table 6a. Items showing DIF between SSI and Non-SSI, (Grade 4, 1992)

Block	Item	Content	Process	Item-type	SSI favored	Non-SSI favored
M3	6. INTERPRET SPEEDOMETER READING	2	2	2	X	
	7. REASON TO FIND AREA OF SQUARE	2	1	1	X	
	12.FIND PROBABILITY	4	1	2	X	
M4	2. COMPARE WEIGHTS	2	1	2		X
	4. IDENTIFY SOLUTION PROCEDURE	1	1	2	X	
	5. INTERPRET PIE CHART DATA	4	3	2		X
M5	6. TOTAL # NEWSPAPERS LEE DELIVERS IN 5 DAYS: MULTIPLICATION	5	1	2	X	
	12. PLACE VALUE- 10S PLACE	1	1	2	X	
	15. IF PATTERN CONTINUES, PUPPY WILL WEIGH 24 POUNDS	5	3	2	X	
	17. GIVEN RECTANGLE DIMENSIONS FIND AREA	2	2	2		X
M6	7. READ A RULER	2	3	1		X
	11. DRAW A GEOMETRIC FIGURE	3	1	1		X
M8	6. SOLVE MULTI-STEP STORY PROBLEM	1	3	2	X	
	13. SOLVE STORY PROBLEM (MULTIPLICATION)	1	2	2	X	
	6. COUNT CUBES IN SOLID	2	1	2		X
	6. COMPARE AREAS OF TWO SHAPES	2	3	1		X
M11	7. DISPLAY DATA ON BAR GRAPH	4	3	1		X
	13. FIND AREA OF FIGURE	2	1	2		X
M12	1. DIVIDE WHOLE NUMBERS	1	2	1	X	
	5. USE MULTIPLICATION TO FIND MISSING NUMBER	1	3	1		X
	6. APPLY CONCEPT OF PARALLEL LINES	3	1	2	X	
	7. ESTIMATE TO SOLVE WORD PROBLEM USING ADDITION	1	3	1	X	
M13	9. OUTLINE SQUARES IN A FIGURE		1			
	10. OUTLINE TRIANGLES IN A FIGURE	3	1	1	X	
M14	1. MULTIPLY A 3-DIGIT NUMBER BY A 1-DIGIT NUMBER	1	2	1	X	
M15	5. SELECT REASONABLE UNIT OF MEASURE	2	1	2		X

Table 6b. *Items showing DIF between SSI and Non-SSI (Grade 4, 1996)*

Block	Item	Content	Process	Item-type	SSI favored	Non-SSI favored
M6	7. READ A RULER	2	3	1		X
M8	8. INTERPRET READING ON A GAUGE	2	2	2		X
	15. VISUALIZE WRITTEN STATEMENT	3	3	2		X
M9	8. RELATE PERIMETER TO SIDE LENGTH	2	3	2	X	
M10	6. COMPARE AREAS OF TWO SHAPES	2	3	1	X	
M11	3. APPLY PROPERTY OF A CUBE	3	1	2	X	
	6. RECOGNIZE EQUIVALENT FRACTIONS	1	1	2		X
	16. COMPARE TO FIND LARGEST NUMBER	1	1	2		X
M12	2. IDENTIFY MULTIPLE OF 5	1	1	2	X	

Table 6c. *DIF and Non-DIF between SSI and Non-SSI (Grade 4, 1992 and 1996)*

1992	1996	# of items
SSI favored DIF	SSI favored DIF	0
	Non-SSI favored DIF	0
	Non-DIF	8
Non-SSI favored DIF	SSI favored DIF	1
	Non-SSI favored DIF	1
	Non-DIF	6
Non-DIF	SSI favored DIF	2
	Non-SSI favored DIF	4
	Non-DIF	75

Table 7  
Grade 8 DIF Items by Content, Process, & Item Type

Table 7a. Items showing DIF between SSI and Non-SSI (Grade 8, 1990)

Block	Item	Content	Process	Item-type	SSI favored	Non-SSI favored
M3	1. ADD FOUR TWO DIGIT NUMBERS (NO CALCULATOR)	1	2	1		X
	3. ESTIMATE LENGTH SHOWN IS TO NEAREST 4TH INCH	2	2	2		X
	7. THE VALUE OF N + # GIVEN N	5	2	1		X
	9. USE CENTIMETER NOT M OR KM FOR PENCIL LENGTH	2	1	2	X	
	10. CHANGE DECIMAL TO A PERCENT	1	2	2		X
	11. 125% OF N IS GREATER THAN N	1	1	2		X
	13. WRITE COMPOSITE NUMBER AS DECIMAL	1	1	1	X	
	15. THE FOURTH FIGURE SHOWN IS NOT A PARALLELOGRAM	3	1	2		X
	16. ONE LITER IN MILILITERS	2	1	2	X	
	17. PERIMETER OF RECTANGLE GIVEN DIMENSIONS	2	2	2	X	
	18. ADD TWO INTEGERS, NEG AND POS	1	2	2		X
	19. AVERAGE AGE	4	2	2		X
	21. DECIMAL BETWEEN TWO OTHERS	1	1	2		X
23. EQUIVALENCE OF ALG EXPRESS	5	2	2	X		
M4	12. SOLVE STORY PROBLEM(FRACT)	1	2	2		X
	14. SOLVE AN INEQUALITY	5	1	2	X	
	15. IDENTIFY COORDINATES ON A GRID	5	1	2	X	
	18. INTERPRET MEASUREMENT TOLERANCE	2	1	2	X	
M5	1. SOLVE A NUMBER SENTENCE	5	1	1	X	
	2. DRAW AN OBTUSE ANGLE	3	3	1		X
	4. APPLY PLACE VALUE	1	1	1	X	
	5. APPLY PART-WHOLE RELATIONSHIP	1	1	1		X
	7. READ A RULER	2	3	1		X
	8. COMPLETE A LETTER PATTERN	5	3	1		X
	9. USE A NUMBER LINE GRAPH	1	1	1		X
	13. EXPLAIN SAMPLING BIAS	4	1	1	X	
	14. GRAPH AN INEQUALITY	5	1	1		X
	15. USE TANGRAMS	3	3	1		X

Table 7a, continued:

M6	7. DRAW A GEOMETRIC FIGURE	3	2	1		X
	9. UNDERSTAND WHEN TO ESTIMATE	1	1	2	X	
	16. SOLVE A PROBABILITY PROBLEM	4	1	2		X
M7	2. SOLVE RATION FOR N	1	2	2		X
	5. GIVEN AVERAGE WGHT FIND COMBINED WGHT	4	1	2	X	
	6. FIGURE A BEST ILLUSTRATES THE STATEMENT	1	1	2	X	
	8. GIVEN NUMBER IN BAG FIND - PROBABILITY DRAW A COLOR	4	3	2	X	
	9. BOX X CUBIC INCHES- MEASUREMENT REPRESENTS VOLUME	2	1	2		X
	15. LENGTH OF RECTANGLE CAN BE EXPRESSED	5	1	2	X	
	18. PROBABILITY SELECT ONE OF TWO TYPES GIVEN NUMBER OF EACH	4	1	2	X	
M8	3. FIND CHECKBOOK BALANCE	1	3	2	X	
	4. SOLVE TWO-STEP STORY PROBLEM	1	3	2	X	
	7. FIND ANGLE IN TRIANGLE	3	2	2	X	
	8. INTERPRET A GIVEN RULE	1	3	2	X	
	9. FIND AN AVERAGE	4	3	1	X	
	12. INTERPRET REPRESENTATION OF FRACTION	1	1	2	X	
	17. ORDER FRACTIONS	1	1	2		X
M9	9. OBJECT IN EARTH WEIGHT WEIGHS ON MOON	1	2	1		X
	13. GIVEN PRICE PER POUND FIND COST OF FRACTIONAL PART	1	2	2	X	
	15. PLOT POINTS ON THE GRID SHOWN	5	2	1	X	
	18. X TO X+10 PERCENT INCREASE IS ?	1	2	2	X	
	20. GIVE RATIO WITH DECIMALS FIND N	1	2	2	X	

Table 7b. Items showing DIF between SSI and Non-SSI (Grade 8, 1992)

Block	Item	Content	Process	Item-type	SSI favored	Non-SSI favored
M3	3. MULTIPLY TWO NEGATIVE INTEGERS	1	2	2		X
	7. READ DIALS ON A METER	2	2	2	X	
M4	3. APPLY TRANSFORMATIONAL GEOMETRY	3	1	2		X
	12. SOLVE STORY PROBLEM (FRACTIONS)	1	2	2		X
	15. IDENTIFY COORDINATES ON A GRID	5	1	2	X	
	18. INTERPRET MEASUREMENT TOLERANCE	2	1	2	X	
M5	16. MARBLE TAKEN FROM BAG - MOST LIKELY TO BE RED	4	1	2		X
	18. GIVEN NUMBER, RATE FAULTY, FIND TOTAL FAULTY	4	3	2	X	
M7	10. IDENTIFY CROSS- SECTIONS OF A CYLINDER	3	1	1		X
M8	7. FIND ANGLE IN TRIANGLE	3	2	2		X
	9. FIND AN AVERAGE	4	3	1	X	
	10. FIND A PROBABILITY	4	3	1	X	
	17. ORDER FRACTIONS	1	1	2	X	
M9	3. SELECT GRAPH FOR INEQUALITY	5	2	2	X	
M10	3. ASSEMBLE PIECES TO FORM SHAPE	3	3	1		X
	4. ASSEMBLE PIECES TO FORM SHAPE	3	3	1	X	
M11	6. RECOGNIZE EQUIVALENT FRACTIONS	1	1	2		X
	7. DISPLAY DATA ON BAR GRAPH	4	3	1	X	
	12. CONVERT HOURS TO ACTUAL TIME	2	3	2		X
	13. FIND AREA OF FIGURE	2	1	2		X
M12	5. FIND COST USING PERCENT	1	2	2		X
M13	10. LOCATE OBJECT ON A GRID	5	3	1		X
M14	3. FIND MEDIAN OF DATA PRESENTED IN GRAPHIC FORM	4	1	2		X
M15	4. READ SCALE	2	2	2		X
	10. UNDERSTAND PLACE VALUE	1	1	2	X	

Table 7c. Items showing DIF between SSI and Non-SSI (Grade 8, 1996)

Block	Item	Content	Process	Item-type	SSI favored	Non-SSI favored
M3	2. WRITE FRACTION THAT REPRESENTS SHADED REGION	1	1	2		X
	3. MULTIPLY TWO NEGATIVE INTEGERS	1	2	2	X	
	9. TRANSLATE WORDS TO SYMBOLS	5	1	2		X
M4	5. INTERPRET PIE CHART DATA	4	3	2	X	
	11. SOLVE STORY PROBLEM (DIVISION)	1	1	2	X	
	14. SOLVE AN INEQUALITY	5	1	2	X	
	15. IDENTIFY COORDINATES ON A GRID	5	1	2	X	
M6	13. EXPLAIN SAMPLING BIAS	4	1	1	X	
M8	3. FIND CHECKBOOK BALANCE	1	3	2	X	
	5. INTERPRET CIRCLE GRAPH	4	2	2	X	
	12. INTERPRET REPRESENTATION OF FRACTION	1	1	2	X	
	17. ORDER FRACTIONS	1	1	2		X
	18. CONVERT TEMPERATURES	5	2	2		X
M10	2. ASSEMBLE PIECES TO FORM SHAPE	3	3	1	X	
	6. COMPARE PERIMETERS OF SHAPES	2	3	1	X	
M11	10. IDENTIFY POINT ON NUMBER LINE	1	1	1	X	
	12. CONVERT HOURS TO ACTUAL TIME	2	3	2		X
M13	10. LOCATE OBJECT ON A GRID	5	3	1	X	
M14	2. IDENTIFY BETTER SURVEY	4	3	1		X
	9. ANALYZE ROAD DETOUR	1	4	3	X	
M15	6. INTERPRET MEANING OF 2X	5	1	2	X	



Table 7d. *DIF and Non-DIF items between SSI and Non-SSI (Grade 8, 1992 and 1996)*

1992	1996	# of items
SSI favored DIF	SSI favored DIF	1
	Non-SSI favored DIF	1
	Non-DIF	7
Non-SSI favored DIF	SSI favored DIF	2
	Non-SSI favored DIF	1
	Non-DIF	6
Non-DIF	SSI favored DIF	10
	Non-SSI favored DIF	3
	Non-DIF	83

Table 8  
*Frequency of DIF Items by Topic and Process*  
*Grade 4 SSI States 1992*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	2	3	2	7
Measurement	1	1		2
Geometry	3			3
Data Analysis, Statistics & Probability	1			1
Algebra and Functions	1		1	2
<b>Total by Process</b>	<b>8</b>	<b>4</b>	<b>3</b>	<b>15</b>

Table 9  
*Frequency of DIF Items by Topic and Process*  
*Grade 4 Non- SSI States 1992*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations			1	1
Measurement	4	1	2	7
Geometry	1			1
Data Analysis, Statistics & Probability			2	2
Algebra and Functions				0
<b>Total by Process</b>	<b>5</b>	<b>1</b>	<b>5</b>	<b>11</b>

Table 10  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 SSI States 1990*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	5	3	3	11
Measurement	3	1		4
Geometry		1		1
Data Analysis, Statistics & Probability	3		2	5
Algebra and Functions	4	2		6
<b>Total by Process</b>	<b>15</b>	<b>7</b>	<b>5</b>	<b>27</b>

Table 11  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 Non- SSI States 1990*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	5	6		11
Measurement	1	1	1	3
Geometry	1	1	2	4
Data Analysis, Statistics & Probability	1	1		2
Algebra and Functions	1	1	1	3
<b>Total by Process</b>	<b>9</b>	<b>10</b>	<b>4</b>	<b>23</b>

Table 12  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 SSI States 1992*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	2			2
Measurement	1	1		2
Geometry			1	1
Data Analysis, Statistics & Probability			4	4
Algebra and Functions	1	1		2
<b>Total by Process</b>	<b>4</b>	<b>2</b>	<b>5</b>	<b>11</b>

Table 13  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 Non- SSI States 1992*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	1	3		4
Measurement	1	1	1	3
Geometry	2	1	1	4
Data Analysis, Statistics & Probability	2			2
Algebra and Functions			1	1
<b>Total by Process</b>	<b>6</b>	<b>5</b>	<b>3</b>	<b>14</b>

Table 14  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 SSI States 1996*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Problem Solv + Ext	Total by Topic
Number & Operations	3	1	1	1	6
Measurement			1		1
Geometry			1		1
Data Analysis, Statistics & Probability	1	1	1		3
Algebra and Functions	3		1		4
<b>Total by Process</b>	<b>7</b>	<b>2</b>	<b>5</b>	<b>1</b>	<b>15</b>

Table 15  
*Frequency of DIF Items by Topic and Process*  
*Grade 8 Non- SSI States 1996*

	Conceptual Understanding	Procedural Knowledge	Problem Solving	Total by Topic
Number & Operations	2			2
Measurement			1	1
Geometry				0
Data Analysis, Statistics & Probability			1	1
Algebra and Functions	1	1		2
<b>Total by Process</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>6</b>

Figure 1. Patterns of DIF by content.

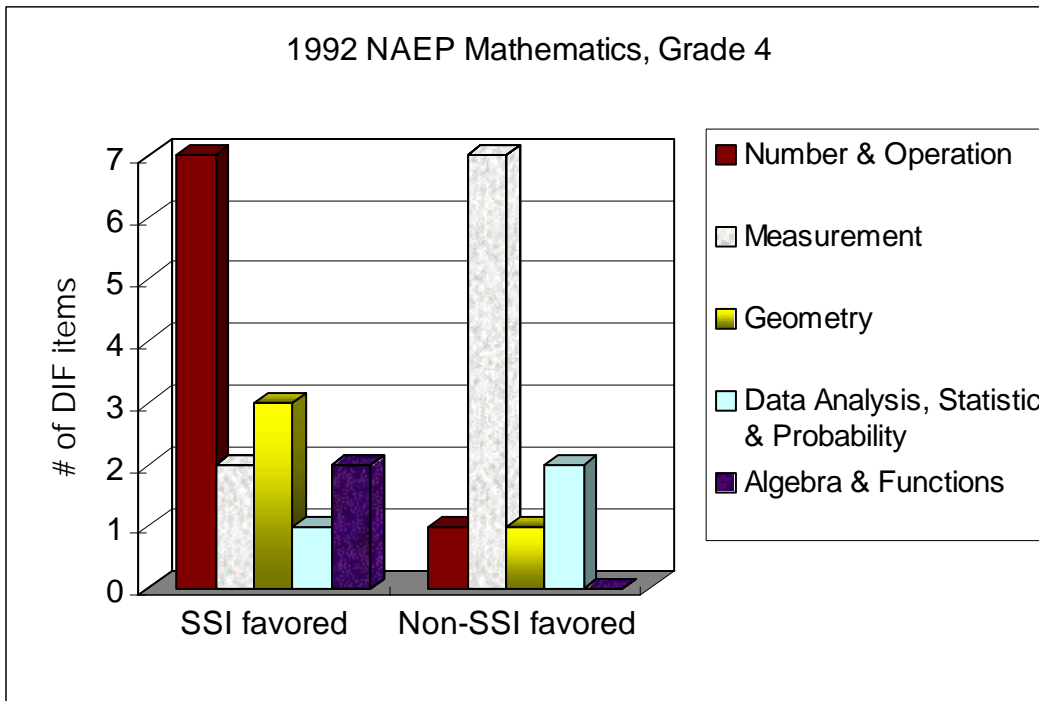


Figure 2. 1996 grade 4 by content.

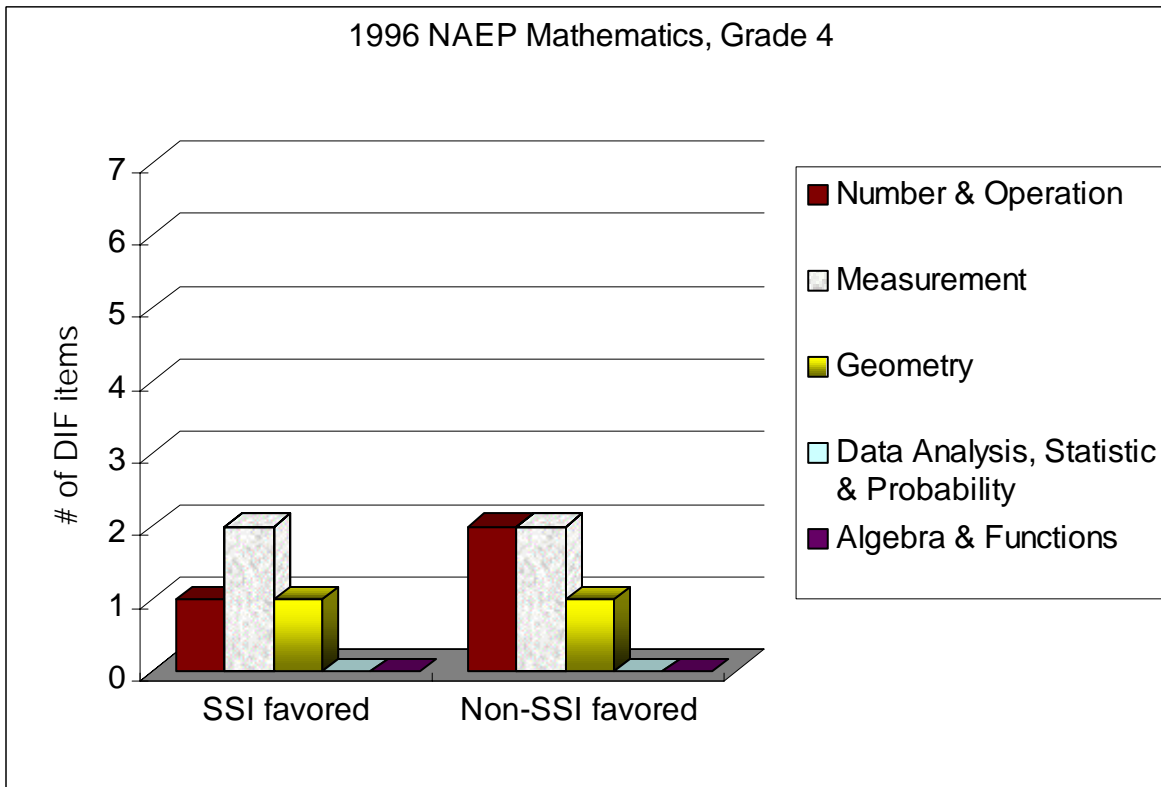


Figure 3. 1992 grade 4 DIF by process.

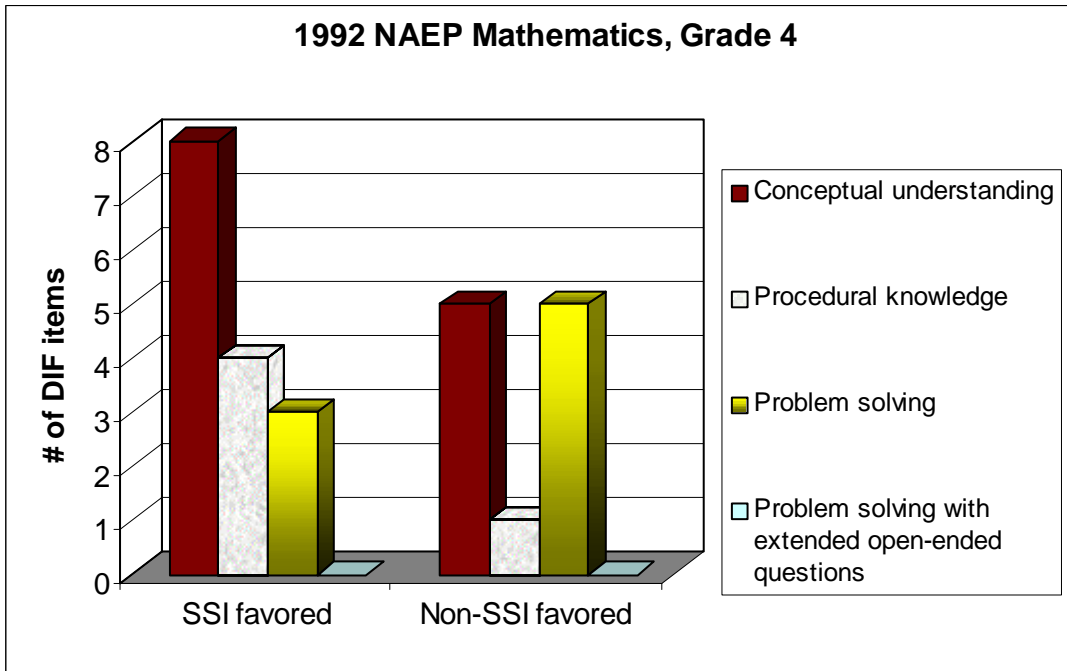


Figure 4. 1996 grade 4 DIF by process.

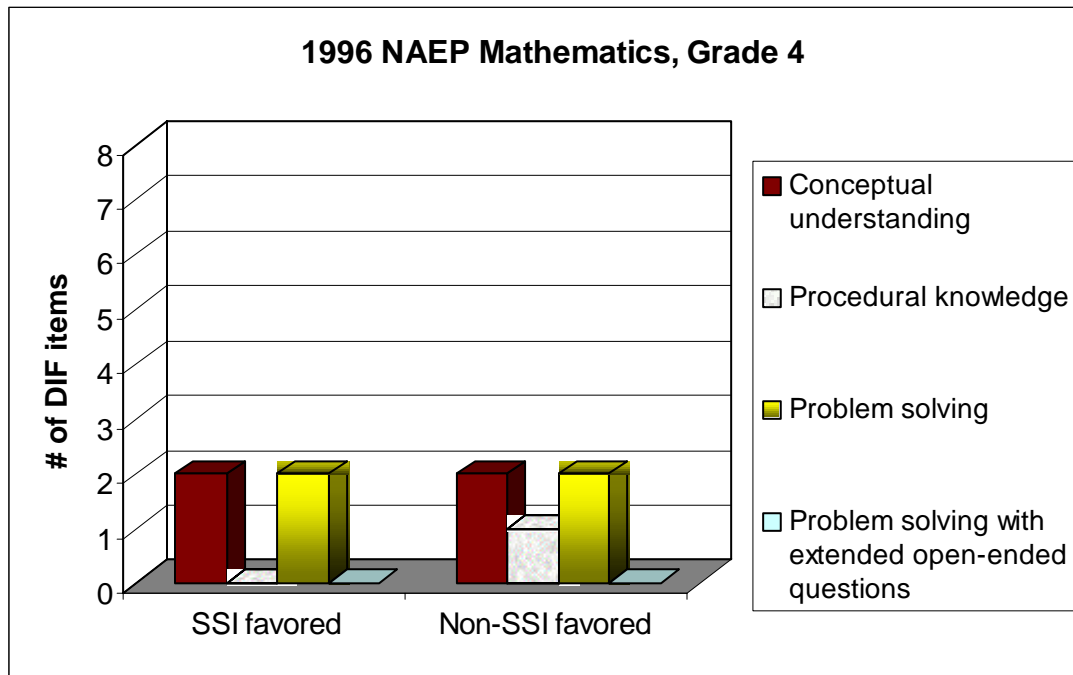


Figure 5. 1990 Grade 8 DIF by Content

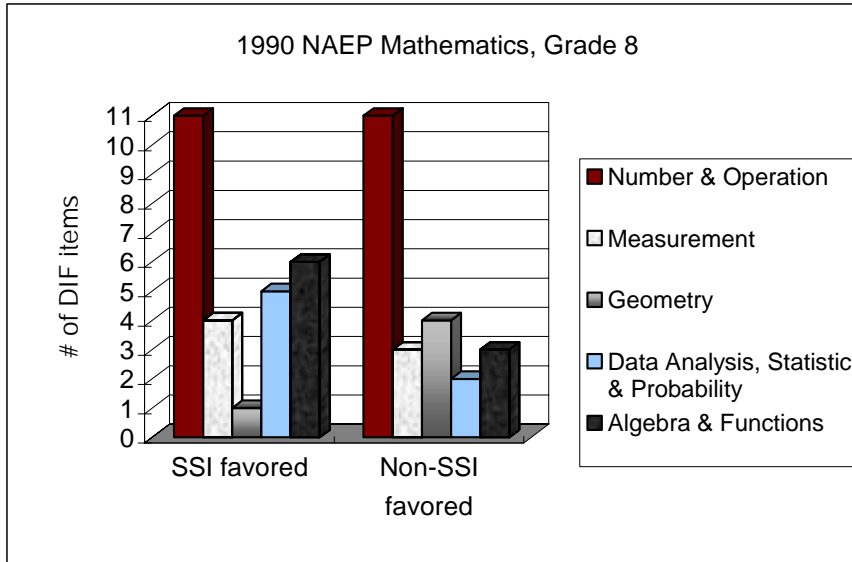


Figure 6. 1992 Grade 8 DIF by Content

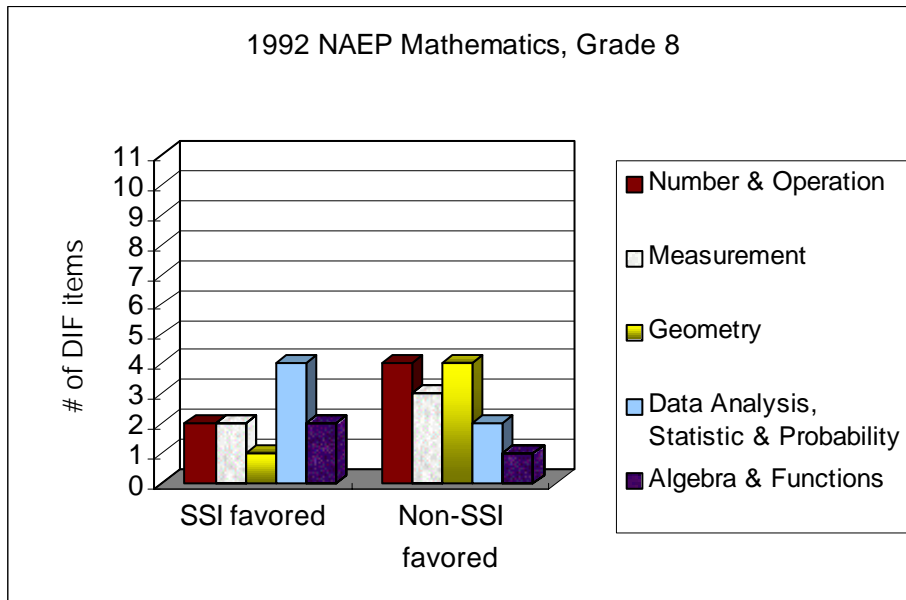




Figure 7. 1996 Grade 8 DIF by Content

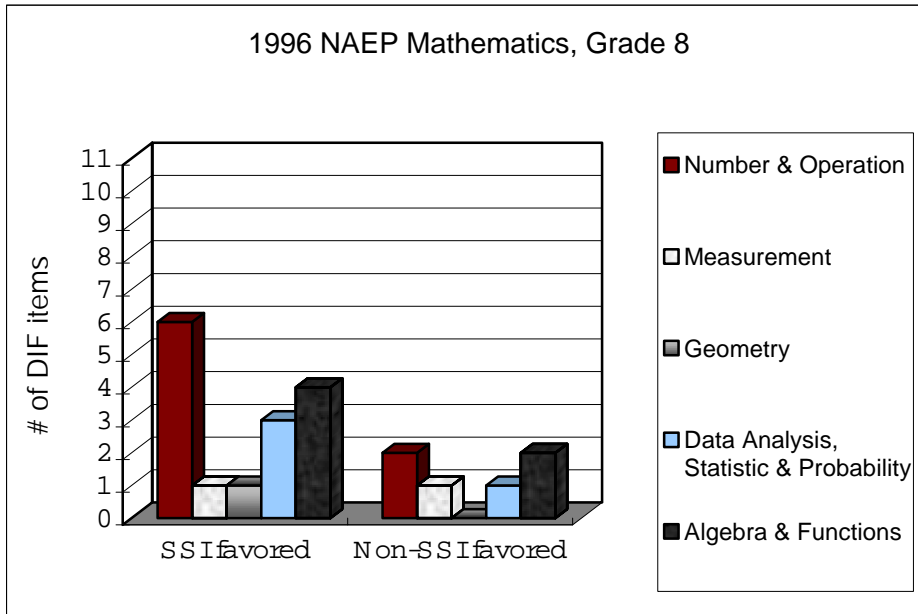


Figure 8. 1990 grade 8 DIF by process.

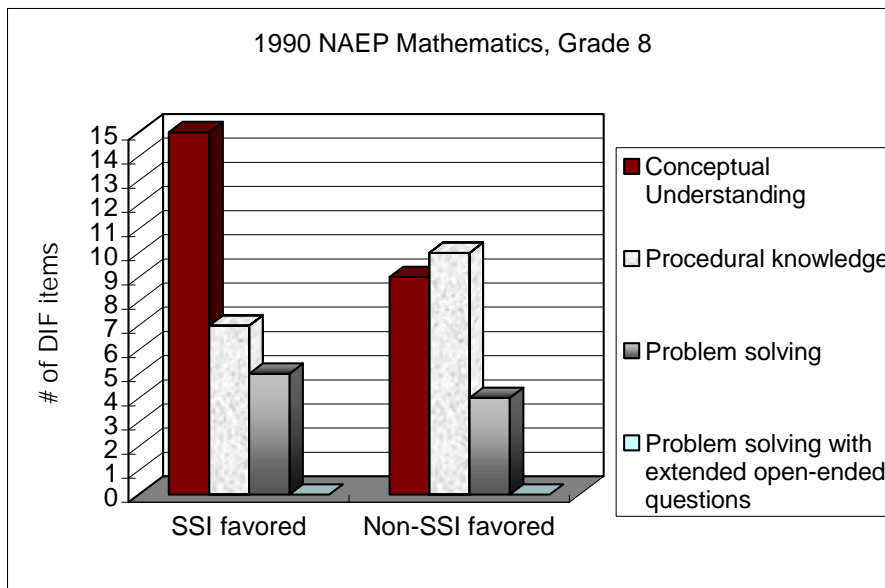


Figure 9. 1992 grade 8 DIF by process.

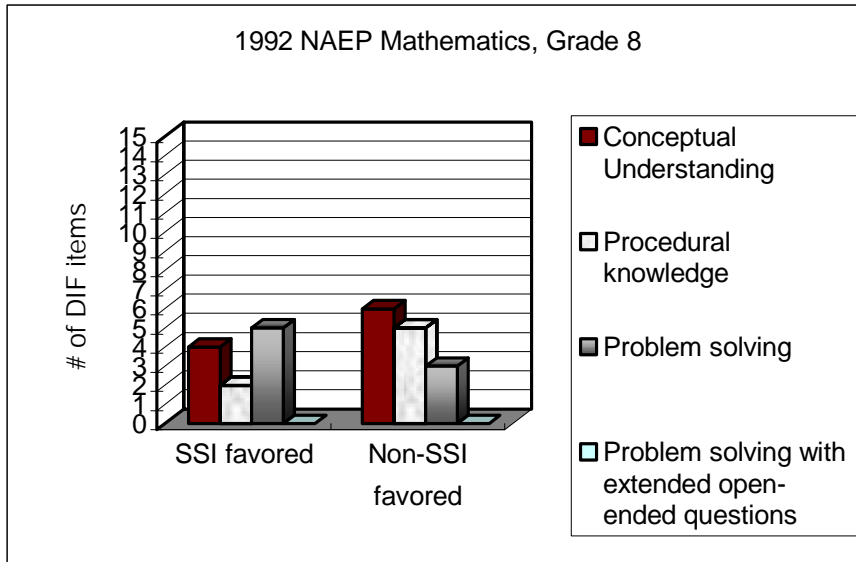


Figure 10. 1996 grade 8 DIF by process.

