

# Performance Trends and Use of Accommodations on a Statewide Assessment:

## Students with Disabilities in the KIRIS On-Demand Assessments from 1992-93 through 1995-96

---

### Maryland / Kentucky Report 3

Published by the National Center on Educational Outcomes

April 1998

---

#### Overview

##### State Assessments Pushing Educational Reform

State assessments and accountability systems are among the principle approaches to educational reform. Despite varying opinions about the approach, policymakers have agreed about the need for these assessments and accountability systems to be inclusive of all children. When exclusion and exemptions from assessments and accountability systems are allowed, numerous problems emerge and questionable practices often occur. For example, both increases in retention rates and rates of referral to special education were documented when exclusion from high stakes assessments was allowed (Allington & McGill-Franzen, 1992; Zlatos, 1994).

Students who receive special education services, in fact, are the most likely students to be excluded from state assessments. And, sometimes, even when they have taken the assessments, these students have been excluded from the accountability systems (e.g., their scores have been excluded when scores are reported, or their scores were not considered when decisions were made about rewards and sanctions). In other words, schools and school districts are not being held accountable for the performance of these students. These policies matter in that as educators consider strategies to improve their standing within an accountability system, students with disabilities and other exempted students can be, and are, ignored in decisions about how resources are re-distributed in personnel and finances.

##### Changes in Special Education Law

The traditional practice of exclusion of students with disabilities from state and district-wide

assessments is no longer possible if states want to continue to receive their special education funds from the federal government. In June, 1997, amendments to the Individuals with Disabilities Education Act (IDEA) were enacted that require, among other things, states to: (a) include students with disabilities in their regular assessments, with accommodations where appropriate, (b) report the number of students with disabilities participating in the regular state assessment, and (c) report the performance of students with disabilities on these assessments in the same way and with the same frequency as they are reported for other students. Furthermore, for those students unable to take the regular assessment, states must develop and implement an alternate assessment by the year 2000, and also report on the performance of students on these assessments in 2000.

Part of the reason for these dramatic requirements in IDEA is the lack of data on the performance of students with disabilities, particularly in relation to educational standards. Individualized Educational Programs (IEPs), which are required for each student receiving special education services, have not provided the information needed to evaluate the performance of students with disabilities. Nor have IEPs provided the information needed to systematically evaluate the programs serving these students, particularly with regard to how well the programs are enabling students with disabilities to progress toward the same standards expected of all students.

### **Kentucky's Inclusive Accountability System**

Despite the overall lack of information on the performance of students with disabilities in relation to state standards, there have been some states that have been striving to include students with disabilities and to gather data on their performance. Chief among these is the Commonwealth of Kentucky. Since 1990, when it incorporated the philosophy that all students can learn, and that the educational system needs to be accountable for the learning of all students, Kentucky has moved toward a totally inclusive assessment and accountability system.

From the time of the establishment of the Kentucky Instructional Results Information System (KIRIS), Kentucky has had a policy of including *all* students. This policy encompassed those students considered to be at risk of failure, and those students with legally identified disabilities, which included both students with formal IEPs and students with 504 plans. To enable the greatest number of students possible to participate in the regular KIRIS assessments, Kentucky established a comprehensive policy on assessment accommodations. Students were allowed to use, during assessment, any accommodation that they used during instruction.

Kentucky's policy of inclusiveness in assessments extended to inclusion of all students in school and district accountability. Thus, the scores of students with disabilities who participated in KIRIS assessments counted in the same way as the scores of students without disabilities.

As a result of its policies, Kentucky has a wealth of information on the performance of students with disabilities on a "regular state assessment." While some researchers have pulled bits of information from this comprehensive set of data, for example, for certain years of testing (Koretz, 1997), no one has yet reported on the complete set of the KIRIS on-demand

assessment data. Particularly lacking are studies describing performance trends over time, an important aspect of the reform-driven system in Kentucky.

The purpose of this report is to take a significant step in that direction—to provide a more comprehensive picture of how students with disabilities are performing *over time* on the KIRIS on-demand assessments. These assessments, which consist of constructed response or essay-like questions over the period 1992-93 through 1995-96, are a significant part of the regular KIRIS assessment, which also includes writing and mathematics portfolio entries. Data from the on-demand assessments reflect the performance of more than 99% of Kentucky's student population in the grades assessed (Grades 4, 8, and 11/12).

It should be noted that Kentucky was unique among states in the early 1990s in that it realized that only a very small percentage of students should be considered unable to participate in the regular assessment, even with the use of accommodations. These students were ones who generally were unable to participate in the general curriculum and who were not pursuing a high school diploma (less than 1% of Kentucky's school population), a determination made by the Admissions and Release Committee (an IEP team). Because Kentucky wanted to be accountable for the learning of all students, it developed an Alternate Portfolio Assessment for these students. Scores from this assessment have an impact equal to those of the assessments of other students in the school accountability process (see Ysseldyke, Thurlow, Erickson, Haigh, Moody, Trimble, & Insko, 1997). (Information on the performance of students on the Alternate Portfolio Assessment is presented in a separate report.)

Thus, students with disabilities in Kentucky were included in KIRIS in one of three ways:

- Participate in the regular KIRIS assessment components in the same way as nearly all other students.
- Participate in the KIRIS Assessment components using accommodations that have been a part of the student's instructional program during the year.
- Participate in the KIRIS Alternate Portfolio Assessment, if eligible according to specific criteria (see Table 1).

As noted previously, the focus of this report is the performance of those students with disabilities who participate in the regular on-demand assessment activities in one of two ways:

- With no accommodations to the assessment administration (because they do not receive instructional accommodations)
- With instructional accommodations applied to the assessment administration (because they receive these accommodations during instruction)

In this report, we also look at the performance of students with disabilities using various types of accommodations.

### **Table 1. Eligibility Criteria for Participation in Kentucky's Alternate Portfolio Assessment**

- a. The student's demonstrated cognitive ability and adaptive behavior itself prevents completing the course of study even with program modifications
- b. The student's current adaptive behavior requires extensive direct instruction in multiple settings to accomplish the application and transfer of skills necessary for functional application in domestic, community living, recreational/leisure, and vocational activities in school, work, home, and community environments.
- c. The student's inability to complete the course of study may not be the result of excessive or extended absences; it may not be primarily the result of visual or auditory disabilities, specific learning disabilities, emotional-behavioral disabilities, social, cultural, or economic differences.
- d. The student is unable to apply or use academic skills at a minimal competency level in natural settings (such as the home, community, or work site) when instructed solely or primarily through school based instruction
- e. For eighth and twelfth grade students with disabilities, the student is unable to
  1. Complete a regular diploma program even with extended school services, schooling, program modifications, and adaptations
  2. Acquire, maintain, generalize skills and demonstrate performance without intensive frequent and individualized community-based instruction

NOTE: These criteria are taken from a May 1992 Kentucky Department of Education Program Advisory. All criteria must be met for a student to be eligible for the Alternate Portfolio Assessment.

---

## Participation Rates and Frequency of Accommodations

Approximately 4% to 10% of Kentucky's students, depending on grade level and year, are identified as having disabilities when participating in the KIRIS assessments. These percentages, which are shown in Table 2 according to grade and year, are based on data coded on the KIRIS answer documents at the time of testing. Because of the nature of this coding and because data are collected in April, it is possible that some students with disabilities are not coded, and thus the percentages of students with disabilities appearing in the table may be different from those reported to the Office of Special Education Programs (OSEP) each year. The December child counts provided to OSEP are by age, while Kentucky's assessment data are dependent on spring grade placement. Thus, direct comparison of numbers is not possible.

Given the available data, we note that the percentage of students with disabilities at the fourth grade level tended to increase some over the four year period, from 8.73% of those assessed using the regular KIRIS assessment components in 1992-93, to 10.18% of those assessed in 1995-96. At the eighth grade level, the percentage remained rather constant, varying between 7.21% and 8.44%. At the 11th/12th grade levels, the percentage of students with disabilities identified as participating in the regular KIRIS increased from 3.71% in 1992-93 to 4.95% in 1995-96.

The use of accommodations during assessments also is documented on the KIRIS answer documents by local district staff. The numbers and percentages of students using accommodations during the 1994-95 and 1995-96 assessments are shown in Table 3. This information was not specifically tagged to students in the data collection process prior to the 1994-95 assessment.

During the two years for which data were available, the percentage of the total population taking the KIRIS assessment components with accommodations remained relatively constant. At the fourth grade level, approximately 8% to 9% of the total student population used accommodations; at the eighth grade level, approximately 5.5% used accommodations; and, at the eleventh/twelfth grade level, approximately 3% of the total student population used accommodations.

Of those students whose tests were coded to indicate that they had a disability, approximately 82% to 84% of fourth graders took the assessment with accommodations during 1994-95 and 1995-96. Approximately 68% to 69% of eighth graders, and approximately 62% at the eleventh/twelfth graders with disabilities used accommodations during these years.

Accommodations used during KIRIS must be consistent with those accommodations that students receive during instruction. For example, if students normally have printed materials read to them during the school year, the printed materials associated with the KIRIS assessments (e.g., constructed response items or portfolio related activities) may be read to these students. If the normal instructional process provides for the students to be able to dictate responses to a *scribe*, such students may dictate responses to items administered for the KIRIS assessments. It is intended that these instructional accommodations be consistent with best instructional practices considering each child's disability and individual needs.

**Table 2. Numbers and Percentages of Students Participating in KIRIS On-Demand Assessments**

<u>Grade/Year</u> <u>Students</u>	<u>Students with Disabilities</u>			<u>General</u>
	Percent	Total Number	Percent	Number
		Number Students		
<b>Grade 4</b>				
1992-93	4322	8.73	45187	91.27
1993-94	4647	9.33	45139	90.67
1994-95	4884	9.88	44541	90.12
1995-96	4864	10.18	42895	89.82
<b>Grade 8</b>				
1992-93	3564	7.21	45852	92.79
1993-94	4383	8.44	47557	91.56
1994-95	4075	7.96	47123	92.04
1995-96	4083	8.05	46609	91.95
<b>Grade 11/12</b>				
1992-93	1336	3.71	34710	96.29
1993-94	1710	4.67	34910	95.33
1994-95	2017	4.88	39286	95.12
1995-96	2080	4.96	39844	95.04
Note: General Students = students without disabilities In 1995-96, 546 students were miscoded. They were not included in either population.				

**Table 3. Numbers and Percentages of Students Using Accommodation During KIRIS On-Demand Assessments**

<b>Grade/Year</b>	<b>Number</b>	<b>Percent<sup>1</sup></b>
<b>Grade 4</b>		
<b>1994-95</b>	4006	8.11
<b>1995-96</b>	4118	8.62
<b>Grade 8</b>		
<b>1994-95</b>	2778	5.43
<b>1995-96</b>	2808	5.54
<b>Grade 11</b>		
<b>1994-95</b>	1242	3.01
<b>1995-96</b>	1288	3.07
<sup>1</sup> Percent refers to the number of students with disabilities using accommodations divided by the total number of students (see Table 1).		

## Trends in the Performance of Students with Disabilities on KIRIS

Kentucky's policy of including students with disabilities in its assessments emphasizes the importance of monitoring the performance of students with disabilities. The overriding purpose of the inclusive policy is to assure that there be continued emphasis on improvement in instructional programs, and that the improvement reaches all components of the student population. Furthermore, the instructional improvement must be reflected in improvement in students' achievement, and it must be shown for all students.

This analysis of student performance is focused on the reading, mathematics, science, and social studies KIRIS on-demand assessments administered during the second accountability cycle (1992-93 through 1995-96). Also administered during this cycle were portfolio assessments in mathematics and writing (as well as the Alternate Portfolio Assessment mentioned previously). Data from the on-demand writing assessment are not included here. During Kentucky's second accountability cycle (ending in 1996), these data were not yet included in accountability calculations. Similarly, data from two other components of the accountability system (arts and humanities; practical living/vocational studies) are scaled differently, and not included in these analyses.

Although KIRIS performance is reported in terms of four performance standards (novice, apprentice, proficient, and distinguished) in each content area assessed, the standards are based on an underlying scale score that is derived from the application of a two parameter graded-response IRT scaling procedure applied to a 7-item test (constructed response items) at the fourth grade and an 8-item test at the eighth and eleventh grades. (Note that in the 1992-93 and 1993-94 school years, the on-demand assessments were administered at the

12th grade; they were moved to 11th grade in 1994-95. Also, the 1992-93 and 1993-94 eighth and twelfth grade assessments were seven open-response items.)

In the base year 1992-93, the scale scores in each of the four content areas had a mean of zero (0) and a standard deviation of one (1). For the following three years (1993-94 through 1995-96), this scale has been equated to reflect the changes in student performance over the 4-year period. The findings are presented here for each grade level.

## Grade 4

In Table 4, we summarize the performance of students with disabilities in terms of the four performance standards (novice, apprentice, proficient, and distinguished) for the four content areas of reading, mathematics, science, and social studies during the 1995-96 testing. Clearly, all students are demonstrating performance at the lower end of the standards (novice and apprentice). While the percentage of students with disabilities in the proficient and distinguished levels is comparable to the percentage of general students in these levels, more students with disabilities than general students tend to fall into the novice category, and more general students than students with disabilities tend to fall into the apprentice category.

In **Figure 1**, we summarize the actual equated scale scores of students with disabilities and general students (i.e., students without disabilities) in Grade 4 on the reading, mathematics, science, and social studies assessments over the four year period from 1992-93 to 1995-96. In general, in this grade the performance gap was narrowed between students with disabilities and the general population of students without disabilities. For example, in reading (where the mean equated scale scores for both groups of students continued to increase over the four year period), students with disabilities began with an equated scale score of  $-.60$ , compared to  $.03$  for general students, while in 1995-96, students with disabilities scored  $.63$ , compared to  $.83$  for general students.

An exception to the tendency to narrow the gap might be in science where, in 1994-95, there was a  $.05$  difference between the two groups as compared to  $.25$  in 1995-96. However, this difference is still smaller than that observed in the 1992-93 school year where the mean for general students exceeded the mean for students with disabilities by  $.65$ .

It is of interest to note that while Grade 4 students in the general population experienced very modest declines in the 1995-96 school year in mathematics, science, and social studies, students with disabilities experienced modest growth in mathematics and social studies. At the same time, the  $.27$  drop in science scores for students with disabilities was the largest noted.

In **Figures 2-5**, we have plotted the mean equated scale scores for each content area summarized in **Figure 1**, along with the points 1 standard deviation above and below these means. These figures include data for Grade 4 students with disabilities, students in the general population without disabilities, and then the total population of students, including both students with and without disabilities. These figures allow for the observation of the overall impact of including students with disabilities in the assessments on total score distributions. The data from which these figures were constructed are included in Appendix A.

**Figures 2-5** demonstrate that the inclusion of test scores from students with disabilities in the

overall score distribution has only a small negative impact. And, as the performance of students with disabilities approached the performance of students in the general population, this impact was even smaller. The data in these figures also indicate that, in general, the gap between the performance of students with disabilities in Grade 4 and the performance of the general population of students is closing.

**Table 4. Numbers and Percentages of Grade 4 Students Scoring within each Performance Standard during 1995-96 Testing**

Content/Standard	Students with Disabilities				General Students	
	Total Population		Percent		Number	
Percent	Number	Number	Percent	Number	Number	Number
<i>Reading</i>						
Novice	593	12	2495	6	49509	7
Apprentice	3303	62	25820	63	49786	63
Proficient	1385	26	12549	30	49425	30
Distinguished	64	1	532	1	47759	1
<i>Mathematics</i>						
Novice	2088	40	11795	29	13883	30
Apprentice	2524	47	23674	58	26198	56
Proficient	320	8	3727	9	4157	9
Distinguished	303	6	2200	5	2503	5
<i>Science</i>						
Novice	2276	44	12194	29	14570	31
Apprentice	2939	55	27758	67	30697	66
Proficient	125	2	1406	3	1531	3
Distinguished	5	0	38	0	43	0
<i>Social Studies</i>						
Novice	1925	37	11949	29	13874	29
Apprentice	2270	53	24254	59	27024	57
Proficient	611	11	5017	112	5628	12
Distinguished	39	1	176	0	215	1

Note: Total percentages may not equal 100 due to rounding

## Grade 8

In Table 5, we summarize the performance of students with disabilities, general students, and the total population in terms of the four performance standards (novice, apprentice, proficient, and distinguished) for reading, mathematics, science, and social studies during the 1995-96 testing. As in Grade 4, most students are demonstrating performance in the lower performance standards (novice and apprentice). Unlike Grade 4, the percentages of students with disabilities in the proficient and distinguished levels are not comparable to the percentage of general students in these levels, except perhaps in science where minimal numbers of students demonstrated performance at the proficient and distinguished levels. Like Grade 4, more students with disabilities than general students performed at the novice category, and more general students than students with disabilities performed at the apprentice level.

In **Figure 6**, we have summarized the performance of Grade 8 students with disabilities and general students over the four years, 1992-93 to 1995-96. It is evident in these figures that the performance gap was narrowed between students with disabilities in Grade 8 and the general population of students without disabilities. For example, in reading (where both groups of students continued to grow over the four year period), students with disabilities began with an equated scale score of  $-1.14$ , compared to  $.02$  for general students, while in 1995-96, students with disabilities scored  $-.44$ , compared to  $.38$  for general students. In other words, the difference in scores of students with disabilities and general students went from 1.15 standard deviation units in the base year to approximately 0.82 standard deviation units in 1995-96.

Students in the general population performed basically the same in 1995-96 as in 1994-95 in reading and mathematics, and experienced declines in the 1995-96 school year in science and social studies. Students with disabilities also experienced a modest decline in science performance, but their performance remained relatively flat in the other content areas. It is also noteworthy that while the gap between the performance of the two groups of students continued to narrow over time, the gap is noticeably larger than that observed at the Grade 4 level. In fact, the smallest gaps at Grade 8 just barely approach the largest gaps at Grade 4.

In **Figures 7-10**, we have plotted the mean equated scale scores for each content area summarized in **Figure 6**, along with the points 1 standard deviation above and below these means. As in the figures for Grade 4 (**Figures 2-5**), these figures also provide for the comparison of the general population of students without disabilities with the total population including both students with and without disabilities. The data from which these figures were constructed are included in Appendix B. Because students with disabilities in Grade 8 have not closed the gap between their performance and that of the students in the general population to the extent that appears to have happened at the Grade 4 level, the impact of the scores of students with disabilities on the total distribution is slightly more noticeable.

**Table 5. Numbers and Percentages of Grade 8 Students Scoring within each Performance Standard during 1995-96 Testing**

Content / Standard	Students with Disabilities				General Students	
	Percent	Total Population Number	Percent	Number	Number	Number
<i>Reading</i>						
Novice	1059	25	1847	4	2906	6
Apprentice	3157	74	37515	82	40672	81
Proficient	50	1	6357	14	6407	13
Distinguished	0	0	40	0	40	0
<i>Mathematics</i>						
Novice	3003	70	15051	33	18054	36
Apprentice	966	23	17022	38	17988	36
Proficient	182	4	7587	17	7769	16
Distinguished	115	3	6099	13	6214	12
<i>Science</i>						
Novice	3631	85	23657	52	27288	55
Apprentice	632	15	21976	48	22608	46
Proficient	3	0	125	0	128	0
Distinguished	0	0	1	0	1	0
<i>Social Studies</i>						
Novice	2920	69	12136	27	15056	30
Apprentice	1258	29	27501	60	28759	58
Proficient	83	2	5705	13	5788	12
Distinguished	5	0	417	1	422	1
Note: Total percentages may not equal 100 due to rounding.						

## Grade 11

In Table 6, we summarize the performance of students with disabilities, general students, and the total population in terms of the four performance standards (novice, apprentice, proficient, and distinguished) for reading, mathematics, science, and social studies during the 1995-96 testing. As for Grades 4 and 8, most students are demonstrating performance at the novice and apprentice levels. Similar to Grade 8, the percentages of students with disabilities in the proficient and distinguished levels are not comparable to the percentage of general students in these levels. And, like both Grades 4 and 8, more students with disabilities than general students performed at the novice level, and more general students than students with

disabilities performed at the apprentice level.

**Figure 11** is a summary of Grade 11 performance on the reading, mathematics, science, and social studies assessments from 1992-93 to 1995-96. Because it is generally accepted that the Grade 11 students taking the KIRIS in the 1992-93 school year were somewhat less motivated than in the following years, performance in the 1992-93 school year at the high school level is difficult to interpret. Therefore, it may be more useful to focus on the three year pattern (1993-94 to 1995-96). As was evident in the Grade 4 and Grade 8 data, students with disabilities in Grade 11 narrowed the gap between their performance and that of the general population of students without disabilities. In reading, the population of students with disabilities obtained an equated scale score of -1.00 in 1993-94, compared to a value of .47 for general students, while in 1995-96, students with disabilities scored -.63, compared to .46 for general students. (If the four year period were considered, the comparable beginning points for the students with disabilities would be -1.38, compared to -.04 for general students.)

The relative distance between performance of students with disabilities is noticeably larger than that observed at the Grade 4 level, and in the 1995-96 school year, somewhat larger than that observed at the Grade 8 level. The performance of both groups remained relatively flat over the 1994-95 and 1995-96 years, with the exception that students with disabilities showed a modest increase in 1995-96, and general students experienced a modest decline in social studies.

**Figures 12–15** are plots of the mean equated scale scores for each content area summarized in **Figure 11**, along with the points 1 standard deviation above and below these means. These figures also provide for the comparison of the general population of students without disabilities with the total population, including both students with and without disabilities. Because the students with disabilities in Grade 11 had not closed the gap between their performance and that of the general population, particularly when compared to changes noted in Grade 4, the impact on the total distribution is slightly more noticeable. The data from which these figures were constructed are included in Appendix C.

#### **Table 6. Numbers and Percentages of Grade 11 Students Scoring within each Performance Standard during 1995-96 Testing**

Content/Standard	Students with Disabilities				General Students	
	Total Population	Number	Percent	Percent	Number	Percent
Percent	Number	Percent	Number	Percent	Number	Percent
<i>Reading</i>						
Novice	1787	81	10869	28	12656	31
Apprentice	390	18	23940	62	24330	59
Proficient	9	0	3810	10	3819	9
Distinguished	0	0	92	0	92	0
<i>Mathematics</i>						
Novice	1671	77	11637	30	13344	32
Apprentice	479	22	17458	46	17937	44
Proficient	24	1	6618	17	6642	16
Distinguished	12	1	2998	8	3010	7
<i>Science</i>						
Novice	1217	56	4541	11	5758	13
Apprentice	962	44	30025	77	30987	76
Proficient	6	0	3786	10	3792	9
Distinguished	1	0	359	1	360	1
<i>Social Studies</i>						
Novice	1714	79	12250	32	13964	34
Apprentice	457	21	21151	55	21608	53
Proficient	15	1	4897	13	4912	12
Distinguished	0	0	413	1	413	1
Note: Total percentages may not equal 100 due to rounding						

## Trends in Performance with Accommodations

The accommodations used by students with disabilities during KIRIS can be grouped into eight categories:

- None. No accommodations; the student with disabilities takes the assessment under the same conditions as all other students.

- Reader/Oral. The assessment is read to the student in a way comparable to the manner in which normal instruction is delivered.
- Scribe/Dictation. The responses to the assessment are dictated by the student and written down by a *scribe* in a way comparable to the manner in which normal instruction is delivered.
- Cueing. The student with disabilities uses mnemonics, problem-solving organizers, semantic organizers, webs, or templates in responding to the assessment in a way consistent with daily instruction and the student's IEP or 504 Plan.
- Paraphrasing. The content of the assessment is paraphrased for the student with disabilities in a way comparable to the manner in which normal instruction is delivered.
- Interpreter. The content of the assessment is signed for the student with disabilities in a way comparable to the manner in which normal instruction is delivered.
- Technological. Technology typically used by the student with disabilities in daily instruction is made available during the administration of the assessment.
- Other. Other accommodations normally made available in the delivery of instruction are made available in the administration of the assessment.

Although the policy permitting use of accommodations has been in place since 1992, collection of data on accommodations used during assessment did not begin until 1994-95. Because students could use those accommodations that they used during instruction, an array of combinations of accommodations was possible. For the analyses presented here, only those accommodations and combinations of accommodations used by a least 100 students were included. The specific accommodations and combinations that met this criterion varied by grade and assessment year, as shown in Table 7.

The most frequently used accommodations were combinations that included paraphrasing and oral presentations. Of the eight primary categories of accommodations, four were not used during the assessment by at least 100 students: cueing, interpreter, technological, and other. The only accommodations or combinations used by at least 100 students across all grades and both years were paraphrasing, and paraphrasing & oral.

## Grade 4

In **Figures 16–19**, we have plotted the 1994-95 and 1995-96 mean equated scale scores (along with the points 1 standard deviation above and below these means) for (a) the total population of students in Grade 4, (b) those Grade 4 students who received any accommodation or combination of accommodations (if used by at least 100 students), separately for each accommodation and combination, and (c) those Grade 4 students who used no accommodations. A separate figure is used for each content area. The data from

which these figures were constructed are included in Appendix D.

Variability in the effects of accommodations is evident in each of the Grade 4 figures. There were two accommodations/combinations in each year that resulted in mean performance above that of the total population. In 1994-95, the oral & dictation combination, and the paraphrasing & oral & other combination, produced mean scores above the total. In 1995-96, dictation, and the paraphrasing & dictation & other combination, produced mean scores above the total. Another observation that is evident in these figures is that some accommodations do not result in higher scores than those obtained by students with disabilities using no accommodations, but in fact, seem to be associated with lower performance. This is particularly evident for the oral accommodation, where in all cases except one, the mean equated scale score is lower for those students using the accommodation than for students with disabilities using no accommodations.

**Table 7. Accommodations and Combinations of Accommodations Included in Analyses (n ≥ 100)**

	Grade 4				Grade 8	
	Grade 11/12		1995-96		1994-95	
	1995-96	1994-95	1995-96	1995-96	1994-95	1994-95
Reader/Oral (Oral)	524	375	597	507	(<100)	161
Scribe/Dictation(Dic)	(<100)	122	(<100)	(<100)	(<100)	(<100)
Paraphrasing (Para)	189	222	308	356	205	270
Oral & Dic	857	907	(<100)	160	(<100)	(<100)
Para & Oral	717	517	1105	1023	553	559
Para & Oral & Dic	(<100)	1620	327	413	(<100)	(<100)
Para & Oral & Other	1344	(<100)	(<100)	(<100)	(<100)	(<100)
Para & Oral & Tech & Dic	(<100)	105	(<100)	(<100)	(<100)	(<100)
None	1007	746	1420	1275	845	792

Note: Entries are numbers of students using the accommodation.

### Grade 8

Mean equated scale scores (along with the points 1 standard deviation above and below these means) for Grade 8 students for each of the content areas are plotted in **Figures 20–23**. Again data are plotted for the total population of students, for each accommodation or combination of accommodations (used by at least 100 students), and for students who used no accommodations during the two testing years. The data from which these figures were constructed are included in Appendix D.

The effectiveness of various accommodations or combinations is much less clear in the Grade

8 data than it was in the Grade 4 data. The highest scoring group of students with disabilities in reading was the one using the paraphrasing & oral combination. However, in 1994-95 this group obtained a mean equated scale score of  $-.23$ , while the total population obtained a mean of  $.29$ . In 1995-96, the corresponding values were  $-.14$  for students with disabilities using this combination, and  $.32$  for the total population. There were no accommodations or combinations of accommodations in either year that resulted in mean performance above that of the total population.

## Grade 11

In **Figures 24–27**, we have plotted the 1994-95 and 1995-96 mean equated scale scores (along with the points 1 standard deviation above and below these means) for (a) the total population of students in Grade 11, (b) those Grade 11 students who received any accommodation or combination of accommodations (if used by at least 100 students), separately for each accommodation and combination, and (c) those Grade 11 students who used no accommodations. A separate figure is used for each content area. The data from which these figures were constructed are included in Appendix D.

At Grade 11, the highest scoring group of students with disabilities in reading was that receiving the paraphrasing & oral combination. In 1994-95, this group obtained a mean equated scale score of  $-.80$  while the total population obtained a mean of  $.42$ . In 1995-96, the corresponding values were  $-.63$  for students with disabilities and  $.37$  for the total population. As in Grade 8, there were no accommodations or combinations of accommodations in either year that resulted in mean performance above that of the total population. In fact, on the equated scale score metric, at both Grades 8 and 11 (where at least 100 cases were reported), the means for students with disabilities, with or without accommodations, were noticeably below the zero point, while the total population (including students with disabilities) is noticeably above the zero point.

## Summary

Given the complexity and variability of findings in the data presented in Figures 16 - 27, it is somewhat difficult to draw out general trends. Clearly, the accommodation or combination that resulted in higher performance varied some depending on the content area, grade, and year. However, the pattern of accommodations being more effective (i.e., resulting in higher scores) in Grade 4 (compared to Grades 8 and 11) seemed to be rather stable within the 1994-95 and 1995-96 data. However, this observation must be qualified in a number of ways because of the differences in the populations of students with disabilities in Grades 4, 8, and 11.

As previously noted (see Table 3), the percent of students with disabilities using accommodations was somewhat different from the elementary to the high school levels. In the 1994-95 school year, 8.11% of the total population was students with disabilities using accommodations, while 9.88% of the total population was students with disabilities overall. Thus, 1.77% of the total population consisted of students with disabilities who used no accommodations during the assessment. In Grade 8, 5.43% of the total population was students with disabilities who used accommodations, while 2.53% of the total population was

students with disabilities not using any accommodations. In Grade 11, 3.01% were students with disabilities using accommodations, while 1.77% were students with disabilities not using any accommodations. Another important consideration when attempting to decipher the findings in Figures 16 - 27 is the decline in the percent of the total population that is identified as students with disabilities as the grade level increases.

---

## Conclusions and Implications

The Kentucky performance data provide one of the first consistent sets of data with a longitudinal perspective and an inclusive approach to assessment and accountability. The data also provide some of the first information on the effects of the use of various accommodations on performance. Overall, these data present a picture of student performance that is not just improving, but perhaps improving at a rate more rapid than that of general education students, at least in certain grades. The data also indicate that while some information is available on the use of accommodations and the differences in performance that seem to be related to their use, there is need for further information on questions related to the effects of accommodation on performance. There is also evidence of the evolving nature of the system within the data. For example, there has been a steady increase in the percentage of students with disabilities assessed using the regular KIRIS from 1992-93 to 1995-96. Yet, the use of accommodations has remained relatively stable.

Also evident are very notable grade level differences, both in the numbers of students with disabilities taking the KIRIS on-demand assessments, and in the percentage of students using accommodations. Thus, we see that approximately 9% of students taking the KIRIS in Grade 4 and 8% of students taking the KIRIS in Grade 8 are students with disabilities, while in Grade 11/12, only about 4% of the students in the assessment are students with disabilities.

One might speculate that as students get older, decisionmakers view the regular assessment as less appropriate for students with disabilities, and therefore, the decision makers are more likely to move students into the alternate assessment system. However, this seems inconsistent with the Alternate Portfolio data. While there was an increase from the initial year (1993) in all grades (4, 8, and 12), the number of students participating in the Alternate Portfolio at the 12th grade for the four year period (1993-1996) has remained relatively stable (189, 241, 201, and 219, respectively, for the four years). Of course, the change in numbers of students with disabilities participating in KIRIS also might reflect the decrease in students receiving special education services at these higher grades, as well as the possibility that students drop out of school when they reach the high school level.

Also of note, and consistent with the findings of Koretz (1997), the percentage of students with disabilities taking the test with accommodations is quite high, ranging from a low of about 62% in Grade 11/12 to a high of about 84% in Grade 4. Although we do not have a good way to estimate the percentages of students with disabilities for whom we would expect that accommodations would be needed, these percentages might be viewed as high. In other states (e.g., Maryland, Rhode Island), the percentage of students using accommodations seems to be lower. Any comparisons of states, however, are complicated by differences in the nature of the assessment, which might play a significant role in determining the need for

accommodations during the assessment. Furthermore, the question of most relevance, and one that can easily get lost when the focus is on data, is an instructional question. When teaching students with disabilities "on grade level" challenging content, and evaluating the student's ability to apply or make use of such content, what instructional accommodations, and in what magnitude, are appropriate?

The performance levels of students with disabilities on KIRIS, regardless of grade level or year of testing, were below those of students in the general population. Using the equated scale scores for comparisons, the performance of students with disabilities was generally considerably below that of other students. The gap between the performance of the two groups of students was more noticeable at the higher grade levels.

Despite the lower performance of students with disabilities, the impact of their lower scores was limited, as suggested by visual inspection of the figures showing performance of the total group, performance of general students, and performance of students with disabilities. Because students with disabilities accounted for less than 10% of the total population, the impact of their scores on the mean scores was marginal. On the Grade 4 reading test in 1993, for example, the mean test score dropped less than one-tenth of a standard deviation unit when the test scores of students with disabilities were merged with the scores of their peers without disabilities. In subsequent years, the change was even smaller.

Of interest, but not all that surprising, was the finding that the spread of the scores grew when the scores of students with disabilities were combined with the scores of their peers without disabilities. Other states and local education agencies are likely to have similar results. Therefore, district superintendents and state departments of education need not be concerned that their student population scores will "look bad" when all students are included in their reports. Thus, state and local summary data can be made complete and inclusive of the total student population.

In all grades, there was a trend of increasing performance over time, at a rate generally exceeding that of the general population of students. The achievement gap between Grade 4 students with disabilities and their peers in the general population decreased dramatically over the four years. In reading, the gap was 0.6 equated scale score units in 1993, and only 0.2 units in 1996. Student growth over the four years was equally as dramatic. In reading, students with disabilities gained an average of 1.2 equated scale score units; in math, approximately 1.0 equated scale score units; approximately 0.9 in science, and 0.75 in social studies. In other words, the gap between students with disabilities and students in the general population seems to be closing over time. This is an extremely important finding because of the concern that educational reforms do not always benefit all students. In fact, Kentucky's reforms were designed with this particular concern in mind. In part, this concern was the one impetus for including all students in the accountability system. The findings suggest that it is working, and doing so even though many other specific instructional approaches have failed to narrow the performance gap between general education and special education students.

It should be noted that the impact of the statewide implementation of an ungraded primary program may also have had an impact on performance and performance trends. This is a matter for further study. Although peers in the general population did not show as much gain, their improvement could also be called dramatic. In reading, the difference between general students who were in Grade 4 in 1996 and those in Grade 4 in 1993 was a positive 0.8 units. In math and science, the difference was a positive 0.5 equated scale score units, and in social

studies, the difference was approximately .25 units. The only drop in mean achievement scores was for students with disabilities on the science test between 1995 and 1996 (-0.25), but over the four year period, growth was still evident.

The data on the effects of specific accommodations are somewhat more difficult to interpret. As noted previously, student performance with accommodations generally is still lower than the performance of general students. Only in 4 of the 104 cases examined (oral & dictation, 1994-95; paraphrasing & dictation & other, 1994-95; dictation, 1995-96; paraphrasing & dictation & other, 1995-96) was the performance of students using accommodations higher than the performance of the total group of students. The fact that paraphrasing and dictation appear in these cases suggests the need for further examination of these accommodations. However, the fact that the unexpected difference in performance occurred so rarely may suggest an anomaly in findings.

It is important to note that it is difficult to draw specific conclusions about the meaning of the higher performance since we do not have the experimental basis that we need to reach certain conclusions. For example, there is no comparison group of students without disabilities receiving the accommodations, a condition that is needed to actually demonstrate that the accommodations may be raising performance in unacceptable ways. Another important comparison might be students with disabilities all instructed with appropriate accommodations, but assessed with and without these accommodations.

The assumption behind the use of accommodations is that they should reveal students' true performance levels, levels that are obstructed from being demonstrated because of the disability. Without the accommodations, we are getting a measure of the effect of the student's disability. We would expect that by circumventing the effect of the disability, performance should be higher. Simply finding higher performance levels does not mean that the accommodation is giving the student an advantage. It may simply be a demonstration that the effects of the disability have been bypassed, and that true levels of performance are evident. Experimental designs are needed to answer these questions.

The finding that in some cases performance with accommodations was lower than performance with no accommodations might suggest that incorrect decisions are being made about the best accommodations for students to use. This may be an incorrect conclusion to reach, however, since we do not know how those students who used the accommodation would have performed without it. Furthermore, the finding may reflect an interaction of decision making with how resources to provide instructional accommodations are distributed across students with different disabilities. For example, only those with the most severe needs may be targeted to receive accommodations. Again, experimental designs are needed to answer these types of questions.

There are additional factors to consider as well. For example, there has been little consideration or study of how the various instructional/assessment accommodations interact with the curriculum at the elementary, middle, and high school levels. In that the curriculum at the elementary level might be thought of as more basic in nature, the accommodations typically applied to both instruction and assessment may be more effective than at upper grade levels where the curriculum might be thought of as more complex in both content considerations and applications. At the elementary level, a student typically is in contact with one or two teachers most of the day across the total curriculum. This might create an environment where it is relatively easy to understand and implement instructional

accommodations consistently across the content. At the middle and high school levels, students typically are engaged with five or more teachers across the curriculum. In this environment, it may be more difficult to coordinate and deliver appropriate instructional accommodations.

For all students with disabilities, the amount of time spent exposed to the daily regular curriculum delivered in a mainstream classroom setting may have an impact on the consistent delivery of instructional accommodations, and therefore, their depth of knowledge in content areas. Another matter of inquiry might focus on an analysis of the actual level of instruction experienced by students with disabilities at the middle and high school levels, as compared to the level or complexity of instruction delivered to the general population of students at the middle school level.

Based on observations resulting from routine interaction with schools as opposed to a well designed random process of inquiry, it appears that some students with disabilities at the middle and high school levels are being instructed at levels below what is normal for the general population. If confirmed to be a systematic problem, this practice would reasonably result in the need for fewer accommodations because the student may be attacking curriculum where such accommodations are not needed. This practice would also result in the population of students with disabilities less likely to have engaged instruction exposing them to the higher content and application expectations typically encountered in the KIRIS assessments.

Certainly these are only several possible hypotheses, and there are numerous others that might be proposed. Obviously, much work remains to be done to understand the performance of students with disabilities in statewide assessments. This study, however, is an important first step in doing so, particularly since it provides a longitudinal look at performance.

---

## References

- Allington, R., & McGill-Franzen, A. (1992). Unintended effects of reform in New York. *Educational Policy*, 6 (4), 397-414.
- Koretz, D. (1997). *The assessment of students with disabilities in Kentucky* (CSE Technical Report 431). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Zlatos, B. (1994). Don't test, don't tell. *The American School Board Journal*, 11, 24-28.
- Ysseldyke, J., Thurlow, M., Erickson, R., Haigh, J., Moody, M., Trimble, S., & Insko, B. (1997). *Reporting school performance in the Maryland and Kentucky accountability systems: What scores mean and how they are used* (State Assessment Series, Maryland/Kentucky Report 2). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
-

## Appendices

### Mean Equated Scale Scores, Standard Deviations, and Counts for Grades 4, 8, and 11: Data Used in Figures 1 through 27

*The Appendices are available when ordering a copy of this report from the NCEO Publications Office.*

NCEO Publications Office  
University of Minnesota  
350 Elliott Hall  
75 East River Road  
Minneapolis, MN 55455  
612 / 624-8561  
Fax: 612 / 624-0879  
email: [scott027@maroon.tc.umn.edu](mailto:scott027@maroon.tc.umn.edu)

---

[NCEO HOME](#) | [SEARCH](#) | [NCEO OVERVIEW](#) | [ONLINE PUBLICATIONS](#) | [INFORMATION EXCHANGE](#)  
[TECHNICAL ASSISTANCE](#) | [PUBLICATIONS LIST](#) | [NCEO RELATED PROJECTS](#) | [RELATED SITES](#)  
[NCEO STAFF](#) | [FEEDBACK](#)

© 1999 by the Regents of the University of Minnesota.  
The University of Minnesota is an equal opportunity educator and employer.  
Home URL: <http://www.coled.umn.edu/NCEO>  
This page was last updated April 6, 1999.