

KERA UPDATE

April 1997

#14

What do these really prove?

Correlations Between KIRIS and Other Tests Commonly Taken by Kentucky Students

Over the past few years, the Kentucky Department of Education has provided only limited information about how KIRIS results compared to other tests. In fact, about the only information released publicly has been a statistical calculation of the "correlation" of scores between the KIRIS test and several other tests. While these correlation numbers are interesting, and do have some value, it is even more important to understand what these numbers may not show.

Take a look at Table 1. This shows that 10 students took 5 different tests. They got identical scores on Test 1 and 2a. Then, student performance uniformly falls on Test 2b, and it falls uniformly again on Test 2c and 2d. The average scores on the test 2 series also fall uniformly, and rather sharply — well over 2 full points from a 5.5 on Test 2a to just 3.1 on Test 2d.

Now, look at the "correlations" for Test 1 as compared to Tests 2a through 2d. Notice there is hardly any change. That is because the best student is still the best, the weakest still the weakest (although there are increasingly more students with a 1 point score on test 2b to 2d).

In fact, for social sciences, these correlation figures (all well above +0.9) would be considered very impressive, even the figure for Test 2d.

Of course, while the correlations remained high, it is very apparent that

overall student performance has fallen dramatically from Test 2a to Test 2d. So, while high correlations mean that the best students do the best on both tests, citing only one-year correlations in a report about these tests hides the fact that actual student performance is falling dramatically on the Test 2 series.

Lets tie that to the KIRIS test and other tests such as the ACT and National Assessment of Education Progress.

Table 2 shows actual correlations that have been computed for KIRIS and

several other tests. Figure 1 shows the best available information on how CTBS scores actually changed in Kentucky for roughly the same time period. Notice the high correlations from CTBS to KIRIS are not reflected in the trends in average CTBS scores.

Table 3 shows actual scores and changes on KIRIS and ACT for 1991-92 and 1993-94. Notice, in particular, that the relatively high one-year correlation between ACT Math and KIRIS Math does not match the trend in the score results over several years. In fact, KIRIS

Table 1

Student	Test 1	Test 2a	Test 2b	Test 2c	Test 2d
1	1	1	1	1	1
2	2	2	1	1	1
3	3	3	2	1	1
4	4	4	3	2	1
5	5	5	4	3	2
6	6	6	5	4	3
7	7	7	6	5	4
8	8	8	7	6	5
9	9	9	8	7	6
10	10	10	9	8	7
Average	5.5	5.5	4.6	3.8	3.1
Correlation to Test 1		1	0.995591	0.98193	0.956541

math average scores improved for 12th Grade students while the ACT scores fell. Thus, **there is a negative correspondence between these tests, not a high positive "correlation"**.

The last table in this Update shows comparable scores on KIRIS and the National Assessment of Educational Progress for 4th Grade reading and math and 8th Grade math. Keep in mind that KERA required a national comparison test in the early days of the reform that would be the same as, or similar to, the NAEP. KIRIS was presented to Kentucky as that national comparison test.

Notice on Table 4, from 1992 to 1994, that NAEP 4th Grade reading dropped a point on the NAEP 500 point scale. At the same time, KIRIS showed a dramatic 17.9 point rise on a 140 point scale. NAEP math results for both 4th and 8th Grade showed a 5 point rise, essentially flat, while KIRIS again showed quite dramatic improvement.

Now, if a straight correlation is taken of the six score sets in this table, the result is just 0.189, a very low correlation. For the reading tests only, the correlation is -1, but with only two points to compare, this really isn't valid. For math only, with 4 score sets available, the correlation is a very low 0.299. Clearly, the correlation of real KIRIS trend scores to real NAEP trend

scores is very low, well below the level needed for high stakes testing. But, because KIRIS is based on changes in scores, these are the correlations that are most revealing and important.

So, citing one-year statistical correlations between KIRIS and other tests may have some value; but, especially for KIRIS, **it is**

potentially very misleading to provide only single year correlations without also including information about test averages and correlations for trends over several years of testing.

It is unfortunate that KDE never publicly released anything but single year correlations for the 1993 and 1994 comparison study of ACT and KIRIS. This is especially unusual because ACT did provide studies for both those years to KDE, and a trend analysis should be possible.

Also, KDE never released a detailed examination of score results for other important tests like the SAT, PSAT, and CTBS. The PSAT in particular is taken by a fair number of Kentucky students, and these results are available for all years of KIRIS testing.

Can concerned citizens and legislators be confident our reform is working if incomplete and potentially misleading statistical processes are provided as the major gauge of progress? **Surely, we can do better!**

Figure 1
Kentucky's Mean CAT-5 NCE Scores From
Kentucky School Advocate, October, 1995
Kentucky School Boards Association

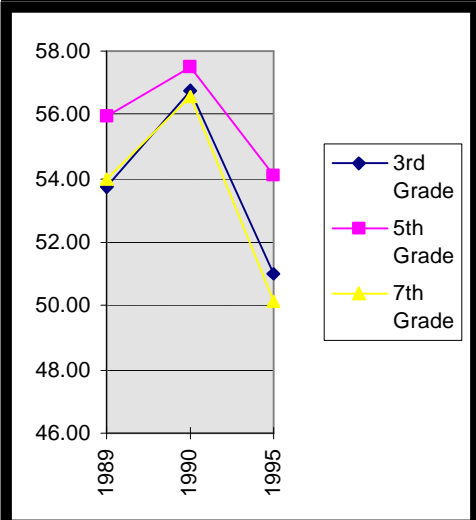


Table 2
Correlations Cited by the Kentucky Department of Education in the August 1994 Kentucky Teacher and reported in KY Office Of Education Accountability's Review of the Measurement Quality of the KY Instructional Results Information System, 1991-1994
(All for 1992-93 School Year)

Grade	NAEP Math	CTBS Math	CTBS Read	ACT Read	ACT Math
4	.72	.77	.86		
8	.63	.86	.63		
12	.81			.59	.75

Table 3
Actual Scores on KIRIS and ACT for 1991-92 and 1993-94 from KY Office Of Education Accountability's Review of the Measurement Quality of the KY Instructional Results Information System, 1991-1994
(Change Scores Include Rounding Error)

Assessment	91-92	93-94	Raw Change	Stand. Change
KIRIS READ	40.3	49.5	+9.2	+0.26
ACT READ	20.5	20.6	+0.11	+0.02
KIRIS MATH	38.2	53.2	+14.9	+0.36
ACT MATH	19.1	19.0	-0.07	-0.02

Table 4
Actual Scores on KIRIS and NAEP for 1992, 1994 and 1996 for Grade 4 Reading and Grade 4 and 8 Math from KIRIS Accountability Cycle 1 Technical Manual KIRIS Cycle 2 Briefing Packet NAEP 1994 Reading Report Card NAEP 1992 Math Report Card

	KIRIS	NAEP
1992 G4 Read	21.9	213
1994 G4 Read	39.8	212
1992 G4 Math	18.1	215
1996 G4 Math	39.0	220
1992 G8 Math	23.8	262
1996 G8 Math	41.0	267