DOCUMENT RESUME

ED 467 509 RC 023 465

AUTHOR	Kifer, Edward
TITLE	Why Research on Science and Mathematics Education in Rural Schools Is Important or the Mean Is the Wrong Message.
PUB DATE	2001-05-00
NOTE	9p.; In: Understanding Achievement in Science and Mathematics in Rural Schools. Conference Proceedings (Lexington, Kentucky, May 21-23, 2001); see RC 023 463.
PUB TYPE	Opinion Papers (120) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; *Educational Research; Elementary Education; *Mathematics Education; Research Needs; Research Problems; *Rural Schools; *School Statistics; *Science Education; Small Classes; Small Schools; *Statistical Distributions
IDENTIFIERS	Variance (Statistical)

ABSTRACT

Educational research, unfortunately, often focuses on finding statistical differences between overall means or averages. Media reports of research routinely present those differences and little else. This paper discusses the importance of considering the spread of the data in addition to the center and how this is relevant to research focused on rural schools. An example from one Kentucky county shows how, in the case of a large group of fourth-grade students with higher average test scores than a much smaller group, the difference between mean scores is misleading and draws attention away from the considerable overlap in the distributions of the two groups' scores. In addition, while the larger group had a higher mean score, it also had many more low-performing students than the smaller group. This county-wide data was broken down further to show the distributions of Group-1 and Group-2 scores in each of the county's six elementary schools. The patterns differed markedly among the schools, suggesting many questions for research. In another example, variance decomposition is used to portray mathematics achievement scores for Japanese and U.S. students in terms of whether the variation is between students, between classrooms, or between schools. Comparisons of achievement status versus achievement growth are also considered. The relevance of these statistical issues to rural education research are discussed in terms of research on small schools, small classes, and the relationships between rural student backgrounds and achievement. (SV)

Why Research on Science and Mathematics Education in Rural Schools is Important Or The Mean is the Wrong Message

Skip Kifer

Edward "Skip" Kifer joined the University of Kentucky faculty in 1972, coming to Lexington from the University of Chicago and the Measurement, Evaluation, and Statistical Analysis (MESA) Specialization. At Chicago, he was a Resource Colleague with a Ford Foundation Teacher Training Program, a member of the College Administration staff, and a statistical consultant to the Department of Education. While associated with the University of Kentucky, he has been a Spencer Foundation Fellow at the University of Stockholm and a visiting professor at both SUNY Buffalo and UCLA. Last year he was the AERA Senior Research Fellow at the National Center of Education Statistics. Dr. Kifer teaches courses in quantitative sequence and introductory evaluation. He played a major role in designing, implementing, and reporting the results of the Second International Mathematics Study and was also one of five persons who planned the Commonwealth of Kentucky's assessment system. His most recent book is *Large-Scale Assessment: Dimensions, Dilemmas, and Policy.*

Why Research on Science and Mathematics Education in Rural Schools is Important Or The Mean is the Wrong Message

Introduction

A recent story in our local paper reported the results of a study suggesting local schools were failing and not living up to the promise of Kentucky's educational reform because there were large differences between the performance of schools with high proportions of poor students and those with low proportions of poor students. This not unusual finding, variously reported as a difference between the average test scores for "rich" schools versus "poor" schools or average differences between "rich" students and "poor" students, is now labeled the achievement gap. There is also an achievement gap between white students and minority students, where it is usually African Americans who are considered the minority.

Educational research, unfortunately, often focuses on finding statistical differences between overall means or averages. Most media reports of results of such research routinely give those differences and little else. Both are committing the cardinal sin of reporting centers of the data without reporting how spread out the data are. They report means and mean differences as though that is all one needs to know in order to understand the findings of the research and what the implications might be for educational practices. **Never a center without a spread** I tell my students and I hope tonight to demonstrate why that is a good axiom and how it might be related to research focused on rural schools.

Some Data

Figure 1(pg. 46) presents some test score results from the Kentucky assessment for 4th grade students from "some "county. The first thing to look at is the table containing the centers. There are two groups, one contains over 1800 students the other over 500. For the larger of the two groups the mean on a scale that goes from 10 to 100 is about 59; the smaller group has a mean of 42. This is an achievement gap of 17 points and would appear to be rather large.

The other parts of Figure 1 show the data so one can get a sense of the spread and distributions of scores. On the left is a box and whiskers plot1 that shows the so-called achievement gap (the middle score for group 1 is higher than the middle score for group 2) but also how the scores overlap. The outliers of Group 2, for example, score at the highest levels. Fifty percent of the Group 2 scores are below 40 but so are about 25% of the Group 1 scores. More than 50% of the Group 1 scores are above 50 but so are more than 25% of the Group 2 scores. The point is that the mean differences can be misleading because otherwise reasonable persons can be lead to believe that average differences mean that *all* persons in one group score higher than *all* of the persons in another group.

The dotplot on the right portrays each of the scores. Notice how much the distributions overlap. But more important, notice that because Group 1 contains so many more students,

Skip Kifer

Edward "Skip" Kifer joined the University of Kentucky faculty in 1972, coming to Lexington from the University of Chicago and the Measurement, Evaluation, and Statistical Analysis (MESA) Specialization. At Chicago, he was a Resource Colleague with a Ford Foundation Teacher Training Program, a member of the College Administration staff, and a statistical consultant to the Department of Education. While associated with the University of Kentucky, he has been a Spencer Foundation Fellow at the University of Stockholm and a visiting professor at both SUNY Buffalo and UCLA. Last year he was the AERA Senior Research Fellow at the National Center of Education Statistics. Dr. Kifer teaches courses in quantitative sequence and introductory evaluation. He played a major role in designing, implementing, and reporting the results of the Second International Mathematics Study and was also one of five persons who planned the Commonwealth of Kentucky's assessment system. His most recent book is *Large-Scale Assessment: Dimensions, Dilemmas, and Policy.*

there are more Group 1 students below the Group 2 mean than there are Group 2 students below the Group 2 mean. In fact, in every part of the distribution one finds more Group 1 than Group 2 students.

The general point I would make about the two pictures in Figure 1 is that if the issue is higher test scores, there is more work to be done in Group 1 than in Group 2. More important, however, is that focusing on mean differences and nothing else is likely to create stereotypes about the groups and make the issue appear to be low performance in Group 2. If there is an issue related to low performance, it is an issue about students not about group averages. And, more students in Group 1 than Group 2 are experiencing the problem.

Looking at Schools

Although the pictures in Figure 1 do a better job of portraying the data, they, too, are limited. Those scores are of students in a county. But students do not attend counties they attend schools. Figure 2 (pg. 47) contains boxplots for six elementary schools in this county. Notice how varied the patterns of differences are. The school represented in the bottom right picture is a school where there are huge differences between the groups. The highest scorers in Group 2 are about at the 50th percentile for Group 1. But look at the boxplots in the upper right of Figure 2. Group 2 scores are higher than Group 1 scores in that picture. The top left picture shows how much less varied the scores for Group 2 re in that school. The middle left picture is interesting because the number of students in Group 2 in that school is so small that there are not enough data to draw the whiskers. Despite their small numbers students in Group 2 have high scores, often higher than the majority of scores of Group 1 in the other schools.

I hope that we have moved beyond the achievement gap of 17 points and to a place where interesting questions can be raised. A first question, of course, is what accounts for these different pictures? Are there policies related to how students are allocated to schools that produce the differences? Do teachers in the different schools treat students in the two Groups differently? Is there some combination of policy and pedagogy, mathematics and science curriculum, that accounts for the differences?

Another set of questions addresses what students experience in the schools. If you were a member of Group 2, which school would you rather attend? Why? If you were a member of Group 1, which school would you rather attend? Why? If the answers to those two questions are not the same, why not?

Another Way to Look at Spreads

Unfortunately my data set does not contain classroom identifications. I would like to look, of course, at each classroom in each school and see what those distributions of scores look like and then start asking questions about the different patterns that I know I would find.

But I do want to talk about classroom differences so I will take another data set and make some slightly different points. Figure 3 (pg. 48) portrays data from the Second International Mathematics Study2 for grade eight students in the United States and grade seven students in Japan.

The pictures are the results of a statistical technique called variance decomposition that seeks to describe, in this case, a set of scores in terms of whether the variation is between students within classrooms, between classrooms with schools, or between schools. The areas of the pie charts are proportional to the total variation in the scores. The pictures allow one to compare the variance components in Japan with those in the United States in terms of what I call status - test scores at one time point, in this case a pretest at the beginning of the school year. A second comparison is of the components of status in the United States versus the components of growth in the United States. Growth is the difference between a posttest at the end of the school year and the pretest.

It should come as a surprise to you that the area of Japan's status pie is larger than the comparable U.S. status pie. (A way to think about this difference is that if test scores were a 100-meter dash the difference between the fastest and slowest runner in Japan is bigger than the difference between the fastest and slowest runner in the United States.) Yes, as the media reports Japan's average score is quite high and among the highest internationally. But, the spread of Japanese scores is among the highest internationally, too. Does that say something about practices in Japanese schools?

The components of the pies (how does one partition the area, the spreads) reflect the structure of schools and schooling in the two systems. Notice that almost all of the variation in Japan is between student differences and there are small differences between schools and classrooms. In the United States the biggest component is between classrooms. This reflects tracking of students into different types of mathematics courses in U.S. schools in the eighth grade. Japan has a common mathematics curriculum for all students. The United States differentiates the curriculum so different students are exposed to different kinds of mathematics. Do these practices lead to different levels of achievement in the two systems? Yes.

I included the growth pie in the United States for a couple of reasons. First, notice that the area of the growth pie is smaller than the area of the status pie. There is less variation to explain when one deals with growth. Second, the components of the growth pie are very different from the components of the status pie. The great majority of the variation in growth is between students; the between classroom component has shrunk substantially.

Reports of mean differences between types of schools or types of students typically are reports of status not growth measures. It can be argued, however, that schools should be judged in terms of their impact on students or the amount of growth that occurs.

But, and this is an important point, the concomitants or correlates of achievement status are different from those of achievement growth. In general, the background characteristics of students are more highly correlated with status than with growth. Effective teaching practices are more highly correlated with growth than with status. Concretely, if one looked at the differences between groups in terms of growth rather than status, those differences would be much smaller for the growth measures. And, if one started to look at the spreads of growth between students and classrooms, those pictures would be very different than one gets with status measures. How to understand the differences between schools and classrooms in terms of growth and spreads is what a researcher should focus on.

And What About Rural Schools

I know this was a long-winded introduction to research with and about rural schools. Yet, it is a necessary prelude because I think those who investigate issues surrounding rural schools are in a position to answer some very pertinent educational questions. And, they will be rewarded if they approach the task in terms of seeking answers to questions about spreads, not centers. These significant questions, I believe, are about small schools, small classrooms, and the relationships among background characteristics of students and their performance in rural schools.

Small Schools

Not all rural schools are small schools. But, I think I am correct in saying that many of the researchers and much of the research about small schools have come from investigators who are interested, too, in rural schools. So I want to ask them to do more research.

I remember reading the Barker and Gump book, Big School, Small School as a graduate student and being convinced then that small schools on the average are better than large schools. Notice, however, that I fell into the centers trap. I think the evidence about small schools, if one thinks about spreads, would suggest that some small schools are better than large schools and others are worse. **A set of research questions about differences among small schools, what makes one small school better than another, and on what important dimensions are they better seems to me to be an interesting set of research questions.** I would like to know, for instance, if a small school is central to a community either geographically, symbolically, or in some other way, does that make it a superior small school. I would like to know how to explain differences in small schools that produce graduates who fare well in say, higher education, compared to graduates who do not fare so well. I would like to know something about the conditions in which teachers work in strong versus weak small schools and how those conditions are related to what teachers do and how students grow. I would like to know about the mathematics and science curriculum in the strong versus weak schools. And, I would like to know something about what teachers do with and about the curriculum. (Note: persons in large schools can ask and try to answer the same questions. I think, however, a first question is how to make large schools smaller.)

Perhaps persons already know the answers to these questions. I know, however, I was surprised by the results of a study of a graduate student in our department who looked at differences between rural schools that did better than expected on the Kentucky assessment versus those who did less well than would be expected. She found that variables such as degrees possessed by the teachers and their grade point averages were not related to the differences between schools. What was related to those differences, however, was the proportion of teachers who attended the school at which they were now teaching. Successful schools had higher proportions of such teachers than did the unsuccessful ones. There was a pattern of these teachers having left their school, gone to a regional university and then returning. Perhaps nepotism is good!

Small Classes

I am under the impression that rural schools (not all of course) are often doubly blessed by being both small and having classes with, relatively speaking, small numbers of students in the classes. This for me is another perfect research opportunity for those interested in rural schools.

The STARS experiment in Tennessee has documented, I believe, the superiority of small class sizes rather than large ones. The research I have read, however, compares the average performance of students who experienced small classes on a variety of variables to those averages for students in larger classes. Again it is a center without a spread. I would like to ask a set of questions about the differences between "good" small classrooms and "not so good" small classrooms. I would be particularly interested in two kinds of outcomes that have been reported to favor small class sizes: 1) the enduring effects of small classes (that is, students from small classes thrive after they leave that environment); and, 2) the smaller average test score differences between minority and majority students who have experienced small classes.

Suppose as a child I were really fortunate and had a really good mathematics or science teacher in a small classroom for my first four years of school. How big a difference would that make as I encounter more mathematics and science in subsequent years? What was good about that good teacher or what was different about that small class, or what was different about the mathematics and science that gave me such an advantage over those who were not in small classes or did not have that good teacher?

Likewise, suppose I was a minority student in a small class with a good teacher. What differences would appear as I continued my schooling? What were the characteristics of the

teacher, the teaching, the content, the curriculum, or the class that made those differences? And, more important, are the answers to my questions about the efficacy of good teachers and small classes the same regardless of the types of students—whether I represent the majority or a minority? If not, why not?

Background Characteristics of Students

This brings me to my third general research issue. **I believe research on rural schools can help us understand better the relationships between backgrounds of students and their performance in schools.** As a corollary, research can inform us about the relationships among performance and student backgrounds between schools. That is, results of the research could paint a clearer picture of the effects of the background characteristics of a student body and the performance of a school. Why do schools with larger proportions of poor students do less well than schools with smaller proportions?

Kentucky has statewide testing that rewards or punishes schools based on whether or not schools increase their test scores. That accountability system imposes unreasonable expectations for more rapid growth for low scoring schools than high scoring schools. Typically the low scoring schools have higher proportions of students receiving free or reduced lunches (the proxy for being poor) than do higher scoring schools.

Periodically one of the educational interest groups in Kentucky trots out a school with large proportions of "poor" students that has high scores in some subject area included in the Kentucky testing program. (The research strategy that collects such results is suspect but I will leave that for another day.) What is interesting is that in most cases it is a rural school that fits the description of having both high scores and high numbers of students on free and reduced lunch. Why is the achievement gap narrower in some rural schools?

I would like to know whether the relationships between poverty and school outcomes are different for rural schools than, say, urban ones. If they are, I would like to know why. Is it because the proxy, free and reduced lunch, for poverty means a different thing in rural areas than urban ones? Is there something about rural schools or their contexts that provide more equal opportunities for students? Is there something about what goes on in rural schools that negates the effects of a student's background on her possibilities for being successful?

If there are differences, I think the answers to such questions are embedded in the spreads of scores of rural schools and classrooms in rural schools, not the centers. What are the characteristics of an effective school or its agenda that differentiates it from a less effective school when, at least superficially, the schools appear to be similar? If a rural school narrows the achievement gap, how does it do it?

Finally, I hope I have raised some interesting questions. I think a consortium like ARSI is the proper arena to begin to answer those questions. There are virtues in collaboration and virtues in looking systematically at important educational questions. Thank you and good luck.

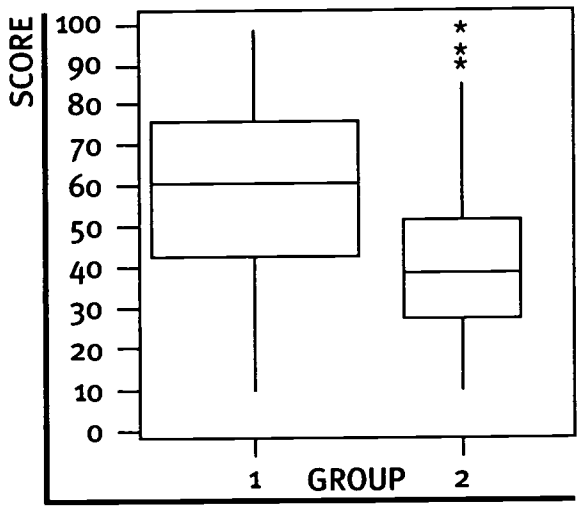
---

1. Boxplots represent the data in the following way: the centerline inside the box is the median or middle score; the top of the box is the 75th percentile and the bottom of the box is the 25th percentile - the box contains 50 percent of the cases. The whiskers cover about 95% of the cases while an asterisk represents outlying or extreme values. The widths of the boxplots are proportional to the size of the samples.
2. The results are similar for TIMSS, the Third International Mathematics and Science Study.

Figure 1.

Scores by group - 4th Grade Students

Variable	Group	N	N*	Mean
Score	1	1809	76	58.8
	2	547	33	41.2



Each dot represents up to 2 observations

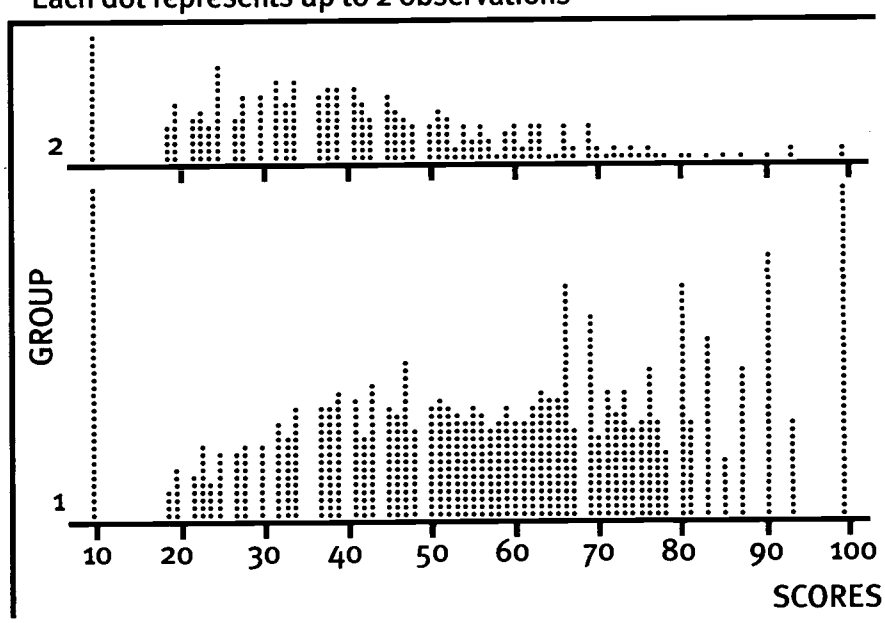


Figure 2.

Within school distributions - 4th Grade Students

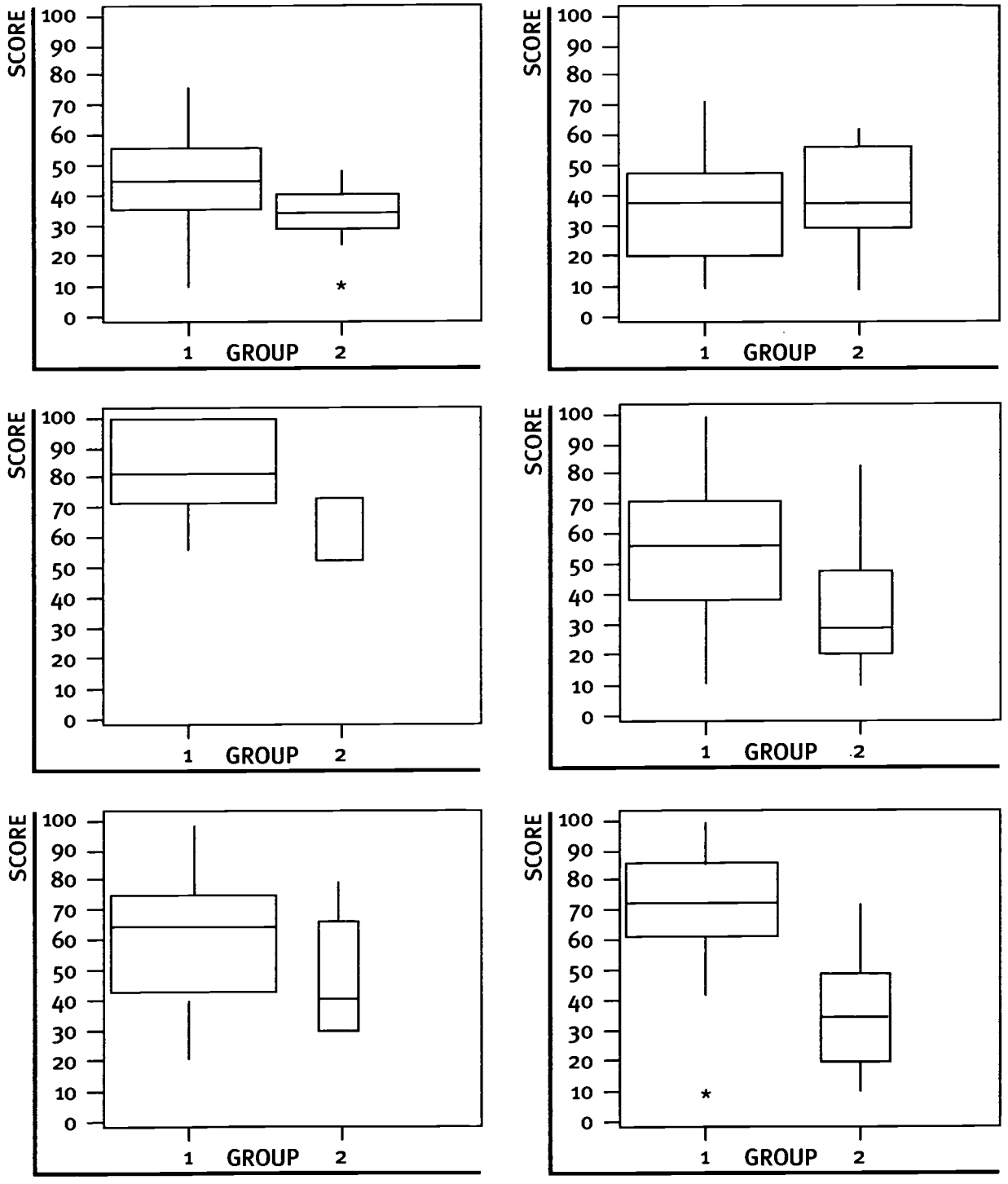
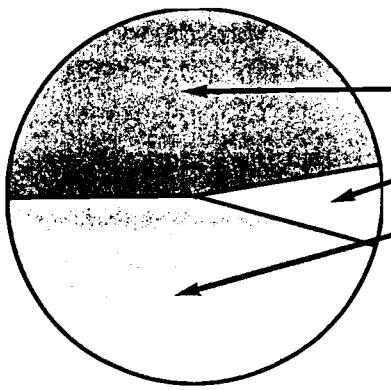


Figure 3.

Variance components of status and growth - 8th grade students

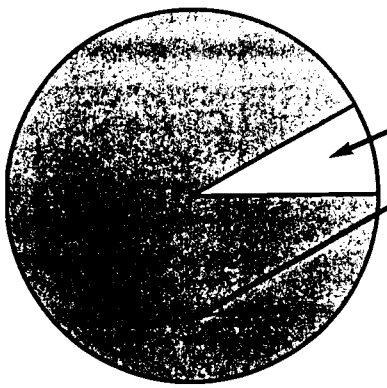


USA - Status

Between Classrooms 48%

Between Schools 7%

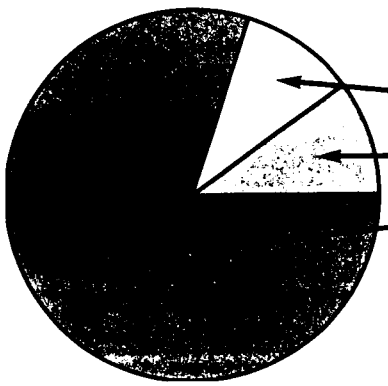
Between Students 46%



Japan - Status

Between Schools and Classrooms 8%

Between Students 98%



USA Growth

Between Classrooms 10%

Between Schools 10%

Between Students 80%

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Why Research on Science and Mathematics Education in Rural Schools is Important or The Mean is the Wrong Message	
Author(s): Edward Kifer	
Corporate Source: Appalachian Rural Systemic Initiative	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: STEPHEN A. HENDERSON ARSI PROJECT DIRECTOR
Organization/Address:	Telephone: (859) 255-3511, X231 FAX: (859) 259-0976
	E-Mail Address: shenderson@arsi.org Date: 8/23/02

(RC 023463)

(over)