Eric Hanushek’s critique of our RAND study was rather one-sided (“Deconstructing RAND,” Check the Facts, Spring 2001). The case for the author’s long-held interpretation that there isn’t “any consistent, positive relationship between increased resources and student learning” is far less airtight than he suggests.

We propose a new hypothesis: Additional resources yield few benefits for nondisadvantaged white students but do raise the achievement of minority and disadvantaged students.

Literature reviews by scholars Alan Krueger, Larry Hedges, and Rob Greenwald have arrived at different conclusions than Hanushek’s concerning the effects of resources on student achievement. A deeper investigation than Hanushek’s also reveals that the “real” increase in resources over time was much less than “threefold.” The author unfairly compares resource increases from 1960 to 1995 to National Assessment of Educational Progress (NAEP) scores from 1971 to 1996. Comparing apples to apples shows a twofold gain in resources from 1971 to 1996. Moreover, Richard Rothstein and Karen Hawley Miles of the Economic Policy Institute have shown that adjusting the inflation rate for the labor-intensive nature of education reduces the real increase to 60 percent. A significant share of this increase, Rothstein and Miles show, went toward special education—the benefits of which would not appear in NAEP data.

In addition, much of this new spending went toward compensatory programs for minority and disadvantaged students, for whom there have been substantial score gains over time.

The aggregate NAEP results cited by the author hide achievement gains of 0.5–0.7 standard deviation for black students and 0.3–0.4 standard deviation for Hispanic students. Disadvantaged white students account for most of the 0.05–0.10 standard-deviation gains of white students generally.

Also supporting my hypothesis is the evidence from class-size reduction experiments in Tennessee and Wisconsin. Both showed significant gains in achievement for students in smaller classes, with larger gains among minority and disadvantaged students. Hanushek, however, suggests that Tennessee’s STAR experiment may have been flawed.

Barbara Nye and her colleagues and Alan Krueger have tested the Tennessee data, searching for evidence of unbalanced attrition out of the control and treatment groups, leakage between the groups, and nonrandom assignment of teachers. They separately found either no flaws or flaws that did not affect the results in any significant way. Despite years of analysis, no one has uncovered evidence that the randomized experiment was flawed.

Hanushek also claims that the class-size results show that reducing class sizes only in kindergarten can raise student achievement. The experiment provided no direct empirical evidence that small classes in kindergarten followed by large classes in succeeding grades would produce sustained effects. However, there is evidence that those students who had only one to two years of small classes in grades K–3 did not sustain their achievement gains through grade 8, while those who had three to four years of small classes were able to sustain their achievement gains.

Our analysis of state NAEP scores also supports our hypothesis. Hanushek raises legitimate methodological concerns about the study. In the end, however, these concerns are empirical issues.

Raising a potential area of concern is not the same as proof that results are biased. There are significant unresolved issues in all studies—about the adequacy of controls for family background, the possibility of differential bias at different levels of aggregation, and the use of weak measures of achievement and spending. Future research must address these unresolved issues in order to ensure more consistency in our experimental and nonexperimental measurements.

Regarding our study, Hanushek raises two issues that are simply wrong. Our sample size was 271 test scores across 44 states, not just 44 as Hanushek implies. Furthermore, estimates from our equations show that modest increases in resources (of $500–$750 per student) can lead to significant score gains (one-third of a standard deviation) among disadvantaged students.

David Grissmer
RAND Corporation
Santa Monica, Calif.
Eric A. Hanushek replies: The most interesting aspect of David Grissmer’s letter is that he makes no effort to defend his own study but refers instead to others’ findings on the relationship between resources and achievement. I take Grissmer’s silence as an acceptance of my conclusions that the RAND study shows a weak relationship between spending and student achievement and cannot be used to identify good policy choices.

The argument that new spending was aimed primarily at minorities and disadvantaged students is similar to other arguments in the RAND study: it bears some relationship to fact but lacks evidence and convincing analysis. It is true that the test-score gap between black and white students narrowed during the 1980s, only to stagnate in the ’90s. However, no evidence shows a parallel shifting of resources to disadvantaged students just in the ’80s. More to the point, the RAND report did not even attempt to analyze racial differences in the way resources affect achievement, even though the NAEP data would have permitted such an analysis. If such differential sensitivity to resources is truly important, RAND’s neglect to analyze racial differences in the way resources affect achievement would bias its results even further.

Grissmer’s reintroduction of highly selective evidence on the effects of extra resources and reductions in class size provides no more support of his work here than it did in the original RAND report.

Stephen Klein et al. surely know better than to claim that separating scores by race adequately adjusts for students’ background. Such stereotyping belies the important heterogeneity within and across each population. To illustrate, the chance that the income of a random black male would exceed that of a random white male approached 40 percent during the mid-1990s. Moreover, the relationship between race and other family characteristics varies from one school district to the next, further complicating Klein et al.’s attempt to reduce the complexity of social life to a single racial dimension.

The researchers claim to have used “all of the schools in Texas” and “all of the 1992–1998 data from NAEP,” as if their inferences come from more than a few averages. Their entire database consists of gains in average scores in math and reading from three specific tests—TAAS, the Texas NAEP, and the national NAEP—for three racial groups. In other words, they have only 27 data points. Such limited data do not provide a sound scientific basis for an informed debate on testing. Had they really used all Texas schools in the second part of their analysis, they would have seen the expected negative relationship between TAAS scores and students’ level of disadvantage—a relationship they did not see with their nonrepresentative sample of 20 schools.

Their testing and peer-review points are, of course, assertions, not facts. If it were known that all of the states’ mandated curricula are well tested by NAEP, it would save considerable money in test development and would quell some intense political debate. Moreover, the quality of the underlying peer-review process should be judged by the outcome.

Correction: The printed version of Figure 1 in Eric Hanushek’s “Deconstructing RAND” (Check the Facts, Spring 2001) contains an error that occurred in the production process and is not to be attributed to the author. A corrected graph, along with its interpretation, appears in the electronic version of the article available at www.edmatters.org and, in more extended form, at www.edmattersmore.org.
Educational McCarthyism

In his article “Cheating to the Test” (Features, Spring 2001), Gregory Cizek refers to allegations about a cheating scandal in New York City, supposedly the largest such scandal ever reported. Cizek cites December 1999 charges by Edward Stancik, the city’s special commissioner of investigation for the public schools, that 52 teachers and administrators had assisted students in cheating on standardized tests.

Former New York City schools chancellor Rudy Crew’s response was to discipline all those named in the report; some were fired, others assigned to desk jobs. All were publicly humiliated and had permanent stains on their reputations.

The United Federation of Teachers subsequently commissioned an independent investigation of Stancik’s allegations by private investigator Thomas Thacher, who previously served as inspector general of the School Construction Authority.

Thacher found that Stancik’s office bypassed basic procedural rights of those accused. The special investigator’s team, Thacher said, pressured children to say their teachers acted wrongly and ignored evidence that pointed toward the educators’ innocence. Accused teachers’ requests to have a union representative present for their interrogations were denied by Stancik’s investigators. Some of the accused were never interviewed at all.

In short, according to the Thacher report, Stancik acted as prosecutor, judge, and jury. The accused teachers and principals had no hearing before an impartial judge to respond to the charges. They were not allowed to confront their accusers or review any of the evidence against them. Thacher concluded that many of those who were named and disciplined were entirely innocent of any cheating; some were reinstated before Thacher’s inquiry.

Stancik also relied heavily on interviews with young children, many of whom were asked to recall testing circumstances two or more years after the events. In some cases, Stancik’s staff never interviewed other adults who were in the classrooms on testing day. Thacher found that many children apparently confused practice tests with the real test.

The Stancik investigation combined some of the excesses of McCarthyism, where the accused were subject to public humiliation without a chance to defend themselves, with the now-discredited practice of relying on children’s memories to destroy the reputations of adult caregivers.

The true scandal was the readiness of public officials and editorialists to condemn these educators without according them a presumption of innocence and a fair hearing.

Diane Ravitch
Brookings Institution
Washington, D.C.

Gregory J. Cizek replies: Diane Ravitch herself acknowledges that it was Rudy Crew, not Edward Stancik, who acted as judge and jury. Furthermore, knowing that it was funded by the New York City teachers union, it is impossible to cast the Thacher report as unbiased. It hypocritically criticizes the Stancik report for being one-sided, yet did not present even one conclusion in support of Stancik’s findings. One has to believe that an unbiased report would have had at least one positive finding.

The Thacher report ultimately reduces to complaints about Stancik’s process, not his findings. I share with Ravitch (and presumably with Thacher and Stancik) an abhorrence for instances of people being railroaded, falsely convicted, or wrongly impugned. We may also agree that, as with any large-scale investigation, better procedures could have been in place, more evidence could have been uncovered, more time spent investigating, and more carefully crafted conclusions proffered. The true scandal is that the message heard by many parents and their children is that cheating is easily excused, argued away on procedural grounds.

Evidence on Integration

Terry Moe finds that inner-city white parents who are opposed to diver-
sity are especially interested in switching to private schools (“Hidden Demand,” Research, Spring 2001). This could lead to the conclusion, says Moe, that inner-city whites “see private schools as a way to avoid integration with minorities.” But Moe is careful to note that a more “benign” interpretation exists: that inner-city whites who choose private schools simply don’t value diversity as much as do whites who see diversity as a strong reason to stay in the public schools.

For some parents, race has been and will continue to be a significant factor in the decision to go private. However, as Moe found and recent research suggests, the prevailing effect of school choice can be to reduce, rather than to increase, racial segregation. This appears to be true in voucher programs targeted primarily at low-income urban families.

Jay Greene, a senior fellow at the Manhattan Institute, found that 19 percent of the students participating in Cleveland’s publicly funded voucher program attended a private school with a racial makeup similar to the Cleveland metropolitan area population, compared with only 5 percent of Cleveland-area public-school students. Fifty percent of voucher students attended racially isolated schools, compared with 61 percent of public-school students in the Cleveland metropolitan area.

In Milwaukee, home to the nation’s oldest and largest voucher program, racial integration is significantly greater in participating private schools than it is in Milwaukee’s public schools. Only 30 percent of students in religious voucher schools attend racially isolated schools, compared with 50 percent of Milwaukee public-school students. This comparative advantage resulted directly from the 1998 expansion of the voucher program to include religious schools, according to my research with George Mitchell.

Howard Fuller
Marquette University
Milwaukee, Wisc.