Letters

Mathematical Ability: Is Sex a Factor?

Benbow and Stanley (12 Dec. 1980, p. 1262) report that, over an 8-year period of studying selected samples of intellectually gifted students, their best male students outperformed their best female students. This is an observation that must be taken seriously because of the large number of volunteers involved and the consistency of the findings. What cannot be concluded from their data, although they speculate to this effect, is that there might be a genetic origin to these sex differences. . . . How then can their consistent findings of male "superiority" be accounted for?

In all the test populations studied by Benbow and Stanley, males outnumbered females. Benbow and Stanley argue, rightly, that we cannot assume the missing females (37 in one year, 418 in another, for example) would have been the highest scoring girls. But we can begin to wonder why fewer girls than boys always volunteered for the gifted programs. If we isolate as a variable "willingness to enter the talent search," we might argue that, although it isn't *likely* that the most talented girls refused to enter the talent search, it is *possible* that, having other options, they did not. Indeed, in research on the same populations, published some years ago, Fox writes (*1*) that it is difficult to get talented girls (and their parents) to agree to participate in a math class for gifted children because of fear of social ostracism. . . .

A second factor might be different out-of-class experiences. In taking mechanical items apart, participating in strategy-memory game-playing, competing in math contests, and playing geometrical or trigonometrical sports (sailing, billiards, and so forth) boys, more than girls, may develop and exercise math-like reasoning powers.

A third consideration is that 7th and 8th graders, however interested they are in mathematics (and the attitude questionnaire administered by the researchers did indicate about equal appreciation for, and interest in, the subject), are experiencing sudden and intense awareness of their adult sex roles and expected behaviors. In addition to the appearance of secondary sexual characteristics at this age (and because of them), 7th and 8th graders are very much aware of the values attached to "masculinity" and "femininity" and, according to many researchers (Fennema, Sherman, Bush), teenagers in general associate mathematics with masculinity (*2*). A question like "Is mathematics a more appropriate activity for boys and men than for girls and women?" might have revealed that there are comparable stereotypes even among the gifted and talented.

Finally, we know precious little about the components of mathematical reasoning and even less about how to test for it.

There is probably no way to conduct a true study of male and female mathematical aptitude so controlled for self-image, out-of-class experiences, and parental reinforcement (or nonreinforcement) that aptitude can be truly sorted out from performance. Moreover, there is no need to do this. If spatial visualization contributes to mathematical reasoning, teach it. Improve math teaching overall, and eliminate all the factors in the culture that discourage children of both sexes and all races from pursuing mathematical study with pleasure and reasonable expectations of success.

CARL TOMIZUKA
Department of Physics,
University of Arizona, Tucson 85721
SHEILA TOBIAS
Overcoming Math Anxiety,
Suite 203, 302 18th Street, NW,
Washington, D.C. 20036

References

1. L. H. Fox, *J. Ed. Gift.* **1** (No. 2), 24 (1978).
2. E. Fennema and J. Sherman, *Women and Mathematics: Research Perspectives for Change* (NIE Papers in Education and Work, No. 8, National Institute of Education, Washington, D.C., 1977); L. Bush, *Girls and Mathematics; The Problem and the Solution* (Apt Associates, Cambridge, Mass., 1980).

Benbow and Stanley "favor the hypothesis" that the sex differences they observed in performance on the Scholastic Aptitude Test in mathematics (SAT-M) "result from superior male mathematical ability, which may in turn be related to greater male ability in spatial tasks." They base this hypothesis, in part, on the contention that "boys and girls have presumably had essentially the same amount of formal training in mathematics," by the 7th grade. There are serious problems with this assumption, as represented by the following findings for elementary school students: girls receive less praise for correct answers than do boys (*1*); boys are praised for participation in academic activities more often than girls (*2*); and teachers sex-stereotype academic fields, making more academic contacts with girls in reading and with boys in math (*3*). From these observations, one would expect males to participate more effectively in academic activities, particularly in mathematics. Equivalence of formal training is therefore not a warranted assumption. . . .

We also note that Benbow and Stanley did not measure spatial ability themselves but cite two published sources for their argument. One of these (*4*) relies on studies now 20 or more years old; the other (*5*) has been superseded by a work (*6*) that clearly voices objection to the male superiority hypothesis.

Certainly this new massive study by Benbow and Stanley reemphasizes the sex differences in mathematical achievement that have been recognized as a serious social problem. Unfortunately, the hypothesis of superior male ability, favored but not substantiated by the authors, received widespread distribution in the popular media (*7*), which did not call attention to the complexity of the problem.

ELIZABETH K. STAGE
ROBERT KARPLUS
Lawrence Hall of Science,
University of California,
Berkeley 94720

References

1. J. Brophy and T. Good, *J. Ed. Psychol.* **61**, 365 (1970).
2. P. Delefes and B. Jackson, *Psychol. Schools* **9**, 119 (1972).
3. G. Leinhardt, A. Seewald, M. Engel, *J. Ed. Psychol.* **71**, 432 (1979).
4. I. MacFarlane-Smith, *Spatial Ability* (Univ. of London Press, London, 1964).
5. J. Sherman, *Psychol. Rev.* **74**, 290 (1967).
6. ——, *Sex-Related Cognitive Differences* (Thomas, Springfield, Ill., 1978).
7. *Time* **116**, 57 (1980); D. Williams and P. King, *Newsweek* **46**, 73 (1980).

. . . The most serious problem with the report by Benbow and Stanley is the underlying presumption that the concept of mathematical ability, as defined by the SAT-M, is theoretically defensible. Such tests sample performance in a domain of learned knowledge and skill: for this reason experts in testing generally recognize the difficulty of separating any measure of ability from achievement, the effects of schooling, and other experience (*1*). In a fundamental sense, we do not yet know what mathematical ability is. . . .

Another suggestion raised by Benbow and Stanley is that girls are particularly deficient in mathematical reasoning ability. The term "mathematical reasoning" also seems to say a great deal more than is justified by the reality of the testing behind it. It refers to performance on word problems. Often, girls—on the average—are reported to do less well on word problems. Before this difference is accepted as real, the possibility of sex bias in the content of word problems should be considered. Analyses of the content of SAT-M (*2*) and other test (*3*) problems have found the content to favor males in a way that can affect problem-solving performance (*4*). Good problem-solvers work with the content of the problem as much as with the mathematical form (*5*), using intuitive understanding of the content to guide choices of mathematical operations. Therefore, familiarity of problem content can make a difference. Girls perform well on tests of computation and algebra where such content bias is not a possibility. . . .

Should we ever discover a genetic and organic basis for mathematical ability, we can be certain at a minimum that some girls will have more ability than almost all boys—a subtlety that keeps getting lost in our "Boys are more or less *X* than girls" language. Why not let performance, with all its pragmatic importance, speak for itself?

SUSAN CHIPMAN
National Institute of Education, Department of Education, Washington, D.C. 20208

References

1. A. Anastasi, in *New Directions for Testing and Measurement*, W. B. Schrader, Ed. (Jossey-Bass, San Francisco, 1980).
2. T. F. Donlon, *Content Factors in Sex Differences in Test Questions* (Research Monograph No. 73-m, Educational Testing Service, Princeton, N.J., 1973).
3. ———, R. B. Ekstrom, M. Lockheed, *Comparing the Sexes on Achievement Items of Varying Content* (ETS Program Report 77-11, Educational Testing Service, Princeton, N.J., 1977).
4. R. G. Graf and J. Riddell, *J. Educ. Res.* **65**, 7 (1972).
5. J. R. Paige and H. Simon, in *Problem Solving: Research, Method and Theory*, B. Kleinmuntz, Ed. (Wiley, New York, 1966).

. . . An underlying fallacy that has been largely responsible for the unwarranted publicity in the popular media which the study by Benbow and Stanley has received is the notion that, if a trait is under genetic control, the expression of that trait is immutable. The genotype of an organism does not determine, in any trivial sense, a single complex phenotype that will be displayed in all environments. One has only to consider something as simple as the height of wheat to realize that a particular genetic strain of wheat will yield different ranges of height in different environments. A determination of the height in one environment will tell one nothing about the height in a substantially different environment. Unfortunately, the history of the past 100 years is replete with examples of deplorable social conditions being attributed to unchangeable innate human differences as an argument for accepting these conditions (*1*). . . . Similarly, medical research into the heritability of diseases would be quite ridiculous if the conclusion were that we should accept as incurable all those diseases which are heritable.

Benbow and Stanley have not shown heritability or innate sex-linked differences in mathematical ability. Even if they had, it would tell us nothing about how to boost the observed performance of girls in mathematics. The real issue is whether or not one wants to see such a change. Attempts to attribute this social difference to a biological difference that would then legitimize the social difference are fallacious at best. . . .

EDWARD EGELMAN
Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254
JOSEPH ALPER
Department of Chemistry, University of Massachusetts, Boston 02125
LILA LEIBOWITZ
Department of Sociology-Anthropology, Northeastern University, Boston, Massachusetts 02115
JONATHAN BECKWITH
Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115
REGINA LEVINE
Department of Zoology, University of New Hampshire, Durham 03824
ANTHONY LEEDS
Department of Anthropology, Boston University, Boston, Massachusetts 02215

References

1. A. Chase, *The Legacy of Malthus: The Social Costs of the New Scientific Racism* (Knopf, New York, 1977).

Benbow and Stanley have made a modest and unexceptionable contribution to one of the real growth industries in contemporary social science. My objections pertain less to content than to editorial form and judgment.

The title of their report is "Sex differences in mathematical ability: Fact or artifact?" Is this appropriate? An "artifact" in science is a spurious result. Whatever legitimate conclusions—social or biological—Benbow and Stanley might have drawn from their data would have been equally "facts." The assumption that truths of the gene are the more profound is thoroughly egregious, but it does not find its way into the text of the report. It is, however, the very stuff of Gina Bari Kolata's commentary (Research News, 12 Dec., p. 1234), which asks the question "Are girls born with less ability?" Yet the data of Benbow and Stanley do not speak to Kolata's question, as Fennema and Gray graciously but fruitlessly warn.

Kolata is scrupulous in seeking out a variety of views. But there is really nothing to talk about. The studies referred to in her article are not germane to her theme, unless one accepts the curious contention (attributed to Benbow and Stanley) that it is "hard to imagine that [the reported differences] are entirely due to socialization." One may, of course, stick closely to the word "entirely" here; but then, a great many once unimaginable things have proved to be true. In any case, the belief that truly great differences in people's characters and situations must have at least some biological roots has been thoroughly aired in other, rather nastier, contexts, and it has not fared very well.

Toward the end of Kolata's article, Benbow advises us that women should accept their differences and stop complaining. At last, one begins to see it: a throng of feminist math students, claiming that society has done them wrong, and seeking federal funds to redress the balance. One of Benbow's colleagues, Tobin, is understandably depressed. She does not "want to think" that she can't "do math like the men do." She speaks not as a scientist, to be sure, but "as a woman," a credential that is, to say the least, widely shared. Even so, her anxiety is enshrined in a special box, near the top of the page.

What does this have to do with science? Nothing. That is the point. As the news weekly of the scientific community, *Science* justifiably tries to offer some features with topical importance and —dare one say it—sex appeal. But it also has, as they say, a reputation to uphold. This is not the way.

DANIEL J. MORAN
Department of History, Stanford University, Stanford, California 94305

Kolata notes that many mathematicians whom we questioned did not know the names of five famous contemporary women mathematicians. Our research (*1–3*) does not support the thesis that girls are born with less math ability than boys. It does suggest that, both in and out of the classroom, environmental influences may be different for females than for males. For example, of the mathematicians whom we questioned, significantly more women than men said that they were discouraged by others in their efforts to become mathematicians, and that they were treated differently, as mathematics students and professionals, because of their sex (*1–2*). To cite another example: on an arithmetical problem-solving test, girls tended to interpret the instructions differently than boys and were more prone to give answers that they thought their teachers wanted (*3*). Such results hint at how difficult it is to control environmental or cultural factors and to ascribe observed differences sheerly or mainly to genetic differences. . . .

For more than a century, sex differences have been reported in mathematical ability, in spatial ability, and in verbal ability, some favoring males and others females (*2, 4*). There is often a large overlap in the distribution of the males and females' scores on the tests, so that there are females with better scores on tests of mathematical ability or spatial ability or with lower scores on tests of verbal ability than males. It is not possible to make predictions from these studies for a particular male or female, nor has a genetic basis been established for any of the differences. Indeed, a genetic hypothesis that has been investigated for two decades—that an X-linked gene determines spatial ability—has been dismissed as unfounded in recent critical reviews (*5*).

The whole issue is reminiscent of the fruitless nature-nurture–IQ controversy, and conclusions about the meaning of differences in scores (between sexes or races) have the same ideological ramifications. Instead of arguing about whether sex differences in mathematical ability are due to heredity or environment, or whether they are "facts" or "artifacts"—particularly, if one means that they are immutable or changeable respectively—it seems more fruitful to study what happens to them under transformation of conditions. In particular, conditions of teaching and learning can be changed. Our studies (*3*) showed that girls responded better to verbal hints in geometric problems than did boys. Using Gestalt psychological principles, we found that spatial configurations could be presented in a structurally clear manner so as to improve their visualization and to eliminate sex differences. Attempts to enhance mathematical learning for both sexes may be challenging ways to channel some of the energy now being used in controversies over the nature and source of sex differences in mathematics.

EDITH H. LUCHINS
Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York 12181
ABRAHAM S. LUCHINS
Department of Psychology, State University of New York, Albany 12220

References

1. E. H. Luchins and A. S. Luchins, in *Women and the Mathematical Mystique*, L. H. Fox, L. Brody, D. Tobin, Eds. (Johns Hopkins Univ. Press, Baltimore, 1980).
2. E. H. Luchins, *Women in Mathematics: Problems of Orientation and Reorientation* (Final Report NSF GY. 11316, National Science Foundation, Washington, D.C., 1976); *Am. Math. Mon.* **86**, 161 (1979); *ibid.*, in press.
3. A. S. Luchins, *Psychol. Monogr.* **54** (1942); ——— and E. H. Luchins, *Rigidity of Behavior* (Univ. of Oregon Press, Eugene, 1959); *J. Genet. Psychol.* **134**, 255 (1979).
4. E. E. Maccoby, and C. N. Jacklin, *The Psychology of Sex Differences* (Stanford Univ. Press, Stanford, Calif., 1974).
5. M. G. McGee, *Psychol. Bull.* **80**, 889 (1979); D. B. Boles, *Child Dev.* **51**, 625 (1980).

The results of Benbow and Stanley on sex differences in mathematical achievement are similar to some results of mine on scientific achievement (*1*). In an international study of the achievement of 14-year-olds in physics, chemistry, and biology, I found that in all 14 countries studied boys scored higher than girls on average. It was also true that within each country the boys' advantage was most marked at the highest levels of achievement. However, the international aspect of the study revealed another dimension, not present in Stanley and Benbow's work. Girls in some countries did better than boys in other countries. When a standard corresponding to the international top 5 percent was defined, this was reached by 11 percent of Japanese girls and 9 percent of Hungarian girls, compared to less than 4 percent of boys in Belgium, the Netherlands, and Italy. It would, I suppose, be possible to argue that there is a genetic difference between children in different countries as well as between girls and boys. But it is equally plausible to suggest that the science education which children receive is responsible for the differences. Some countries teach science more efficiently than others; but in all the countries studied science has a masculine image and is taught in a way which is oriented toward boys rather than girls (*2*).

ALISON KELLY
Department of Sociology, University of Manchester, Manchester M139PL, England

References and Notes

1. A. Kelly, *Girls and Science: International Study of Sex Differences in School Science Achievement* (Almqvist & Wiksell, Stockholm, 1978).
2. These ideas are explored in *The Missing Half: Girls and Science Education*, A. Kelly, Ed. (Manchester Univ. Press, Manchester, 1981).

So little of our report is quoted directly that it seems desirable to reproduce its concluding paragraph: "We favor the hypothesis that sex differences in achievement in and attitude toward mathematics result from superior male mathematical ability, which may in turn be related to greater male ability in spatial tasks. This male superiority is probably an expression of a combination of both endogenous and exogenous variables. We recognize, however, that our data are consistent with numerous alternative hypotheses. Nonetheless, the hypothesis of differential course-taking was not supported. It also seems likely that putting one's faith in boy-versus-girl socialization processes as the only permissible explanation of the sex difference in mathematics is premature."

In this context, "superior male mathematical ability" and "greater male ability in spatial tasks" mean only that boys tend to score higher on the SAT-M and spatial tests (*1*) than girls do. Unfortunately, many readers (such as Stage and Karplus) interpret "superior ability" as meaning inherently, intrinsically, or genetically abler. As national norms show, girls tend to score lower on spatial tests, such as the Differential Aptitude Test (*1*), whether or not for environmental reasons.

We postulated that "endogenous" variables (*2*) *may* be involved. Endogenous sex differences have been documented in a wide variety of organisms, including humans (*3*). We have carried out more research and helped more able young girls mathematically (*4*) than most other investigators in the country. We believe that the last sentence of our brief report is consistent with the present state of knowledge.

Because the girls in our studies were bright, eager, volunteer participants in a mathematics talent search and were matched with boys on in-school mathematics tests, most of the proffered explanations that the sex differences are defects of the SAT-M as a measure of mathematical reasoning ability are irrelevant. The SAT is designed mainly for above-average 11th and 12th graders, 4 or 5 years older and more advanced in grade placement than participants in the Study of Mathematically Precocious Youth's (SMPY's) annual search for youths who reason extremely well mathematically. Thus, the SAT-M almost certainly functions far more at the analytical reasoning level for the SMPY testees, who have not had many formal opportunities to develop their abilities, than it does for high school juniors and seniors who have already studied rather abstract mathematics for several years.

Stage and Karplus challenge our assumption that our males and females received similar *formal* training in mathematics. We doubt that the studies they cite are relevant to the mathematical reasoning ability of our subjects, who are highly able and well motivated (*5*). In addition, Stage and Karplus object to our references relating sex differences in mathematical reasoning ability to spatial ability, saying that one was "20 or more years old" and the other had been "superseded." To their first objection we reply that research findings do not automatically become invalid with age. Their second objection has been in turn "superseded" by our own study of the spatial ability of a subset of SMPY examinees, as well as by other studies (*6*).

The male-female ratio of participants tended to be remarkably stable from year to year at about 57 to 43. We postulate that fewer girls could meet the qualifying score for the search, in addition to the fact that more boys than girls found entering a mathematics competition appealing. In eight talent searches thus far, involving about 34,000 participants, we have found no evidence that the ablest girls tend not to enter. Moreover, the January 1980 and 1981 searches involving 24,000 participants were based on verbal as well as mathematical qualifications, and the sex ratio of participants was 1 to 1. Nonetheless, the mean difference between sexes on the SAT-M was exactly the same (32 points) as in the January 1979 search, when one could qualify only on a mathematics test. Also, even if we accept a 1.3 boys to 1 girl bias in participation, that could not account for the great difference between boys and girls in *high scores* earned.

Egelman *et al.* erect a straw man we do not recognize from our report. For our groups, Chipman also seems to be grasping at straws, however relevant the ETS research she cites may be to 11th and 12th graders who take the SAT. The "artifact" in the title of our article, to which Moran refers, indeed means a spurious result—specifically, the frequently seen assertion that girls do mathematics as well as boys until they drop out of the courses. In addition, however, the primary definition of the word "artifact" is "any object made or modified by man" (*Random House Dictionary of the English Language*, College Edition). Surely, it is reasonable to ask whether the differences are man-made or endogenous. Kelly's observations about physics, chemistry, and biology may be more relevant to learning mathematical concepts and computation than to mathematical reasoning ability. Kelly has not established causal links between male images and the sex difference in ability.

The statement in the 16 January 1981 *Science* editorial by Schafer and Gray (p. 231) that "Not a single student identified by [SMPY] as mathematically precocious—boy or girl—has gone on to do graduate work in mathematics. . . ." is simply incorrect. For example, an 18-year-old is a 4th-year graduate student in "pure" mathematics at Princeton University. Two more are at the Massachusetts Institute of Technology. Three are, or were, at Berkeley, Stanford, and Johns Hopkins. As graduate students, SMPY's protégés have not yet been studied systematically. We know of only 19, a majority of whom major in mathematics, computer science, electrical engineering, physics, and other fields in which high mathematical ability is important.

We deeply regret that press coverage of our brief report confused the issues, rather than alerting people to the *magnitude* of the sex difference. The situation is far worse than most persons realize. For example, of the 7500 boys and 7500 girls in our current talent search (selected for overall intellectual ability), 23 boys and no girls scored 700 or more on the SAT-M. Our search nationwide found another 19 boys but no girls scoring 700 or more. Let's face these dismaying findings squarely and search hard for causes, whatever they may be. Our data clearly show a large, important sex difference *before* the well-documented and intensive socialization processes operating during puberty. Moreover, our data show that extensive boy-versus-girl socialization processes during this period seem to have little, if any, effect on mathematical reasoning ability. At the very least, one must discount differential course-taking. It would also seem prudent not to rule out endogenous explanations for these sex differences entirely.

It is not the method of science (or *Science*) to ignore published facts or provide a forum for subjective judgments and anecdotal evidence.

CAMILLA PERSSON BENBOW
JULIAN C. STANLEY
Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218

References and Notes

1. G. Bennett, H. Seashore, A. Wesman, *Manual for the Differential Aptitude Test* (Psychological Corporation, New York, ed. 5, 1974), forms S and T.
2. The word "endogenous" has a much broader meaning than "genetic"; the latter word does not appear in the report.
3. R. Goy and B. McEwen, *Sexual Differentiation of the Brain* (MIT Press, Cambridge, Mass., 1980); J. Levy, *The Sciences* **21** (No. 3), 20 (1981).
4. Thus far, the Johns Hopkins groups have reported their rationale and findings in seven books and more than 100 articles, including the following volumes published by the Johns Hopkins University Press: *Mathematical Talent* (1974), *Intellectual Talent* (1976), *The Gifted and the Creative* (1977), *Educating the Gifted* (1979), and *Women and the Mathematical Mystique* (1980).
5. C. Benbow and J. Stanley, in preparation.
6. C. Benbow, M. Kirk, L. Daggett, J. Stanley, in preparation; S. Burnett, D. Lane, L. Dratt, *Intelligence* **3**, 345 (1979); S. Cohn, in preparation.

Substance	Lethal and detrimental cells (%)	Number of progeny examined
Cerophyl	1.01 ± 0.21	1568
Glass beads + uninduced S-9	1.36 ± 0.33	1904
Glass beads + induced S-9	1.41 ± 0.42	1888
Benzo[*a*]pyrene + uninduced S-9	3.2 ± 0.39	1440
Fly ash + uninduced S-9	3.7 ± 0.69	2992
Heated fly ash + induced S-9	3.9 ± 1.6	1728
Heated fly ash + uninduced S-9	4.6 ± 1.3	1280
Fly ash + induced S-9	9.3 ± 1.7	1296
Benzo[*a*]pyrene + induced S-9	12.5 ± 5.8	1664

Errare humanum est. In the erratum published in the issue of 13 February (p. 656), Tables 1 and 2 from the report "Mutagenicity of fly ash particles in *Paramecium*" by J. Smith-Sonneborn *et al.* (9 Jan., p. 180) were printed incorrectly. Significance lines missing in the original tables were also missing in the erratum. With determination, we again reprint the tables as they should have appeared (Table 1 above, Table 2 at right).

Substance	Lethal cells (%)	Number of progeny examined
Glass beads	3.14 ± 0.33	624
HCl-extracted particles	4.77 ± 1.25	960
DMSO-extracted particles	8.21 ± 2.00	656
Unextracted particles	14.09 ± 5.29	896