GENDER AND MATH PERFORMANCE: DOES BIOLOGY HAVE IMPLICATIONS FOR EDUCATIONAL POLICY?

Jon Beckwith
Harvard Medical School, Boston, Massachusetts

The use of new biological theories of human behavior to inform educational policy is criticized. Authors who suggest that detecting a biological basis for a particular human ability indicates the unchangeable nature of that ability are committing the fallacy of biological determinism. One study is chosen to illustrate the problems inherent in attempts to link a biological substrate with human achievement. The study by Benbow and Stanley on differential math performance between boys and girls is analyzed in terms of the assumptions which underlie it. An unstated assumption of the Benbow-Stanley study is that socialization factors, in particular those that cannot be quantified, are not important in the development of math test performance. The authors also commit the error of reification, attributing to a test score a physical meaning. Evidence is described indicating the strong influence of environmental factors on SAT, IQ, and other tests. Finally, the appearance of this study and the publicity it received are seen as a reaction to the emergence of the women's movement in this country in the last 15 years.

In recent years, a rash of articles has appeared suggesting that new information from the biological sciences holds important lessons for educators. An entire issue of *The Massachusetts Teacher* entitled "The Brain—Education's Next Frontier" was devoted to the implications for educational policy of new research findings on the brain. Differential brain lateralization between boys and girls requires different modes of teaching math for the two sexes, according to one article (Loviglio, 1981). Plateaus in brain development in junior high school students require modifying curricula at this stage, according to another (Burbank, 1981). These are only the latest in a long history of biological theories which have been used to influence educational practices. While these theories have consistently been shown to lack a solid scientific basis, they have still had considerable impact.

The idea that brain structure, our genes, or other biological factors constrain our educational development is not new. Louis Agassiz, a prominent Harvard zoologist, proposed early in the 19th century that the brain casing of blacks was smaller than that of whites (Gould, 1977, p. 438). He claimed that too much education of blacks would expand the brain size beyond the capacity of its casing, causing serious brain damage or death. Later in that century, craniologists such as Paul Broca proposed that women would never reach the intellectual heights of men, since their

brains were smaller. Gustave Lebon, a founder of social psychology and one of Broca's school, suggested that "A desire to give them [women] the same education, and as a consequence, to produce the same goals for them, is a dangerous chimera" (Gould, 1981, p. 104).

In this century, Cyril Burt, a dominant figure in educational psychology who believed that the differences in intelligence between rich and poor were genetic, played a major role in the introduction of the 11-plus system in England (Hearnshaw, 1979). In this country, psychologist Arthur Jensen proposed in an article in the *Harvard Educational Review* in 1969 that blacks' lower achievement in society was due to their inferior genes (Jensen, 1969). His suggestion that compensatory education programs for blacks would not work was used by some in government as an argument for ending such programs.

Most recently, two psychologists from Johns Hopkins University, Camilla Benbow and Julian Stanley, published an article in *Science* in which they purported to show that males were superior in math ability to females (Benbow & Stanley, 1980). The impression was given to the public that a genetic basis for male-female differences in math performance had been found. Benbow reinforced this impression when she stated, in an interview, that women would be "better off accepting their differences," implying that they were innate and, therefore, unchangeable (Kolata, 1980).

All of these arguments, including this latest example, share a number of common features:

1. They propose a biological basis for a socially observed difference.
2. They imply that if a difference is biological it is unchangeable.
3. They commit the error of reification. They assign to a socially constructed trait a physical attribute.
4. They support the idea that the dominant sectors of society (male or white) are superior.
5. They come at times in history when society is convulsed over questions of sex or race.
6. They are scientifically invalid.

Finally, they are all examples of a reductionist approach to studying such problems. They attempt to reduce the causes of complex social phenomena to simple units—ultimately genes. As Lionel Tiger, one who has taken this approach, has put it, "If male dominance extends over the whole species and has existed for so long, we seem constrained by the law of parsimony to look first into the biological information and theory at our disposal for an explanation" (Tiger, 1970 p. 35). But there is no "law of parsimony." It is simply being used here as a smokescreen for ignoring the complexity of social and cultural evolution in favor of simplistic biological determinist explanations.

In this paper I will focus on Benbow and Stanley's study of differential math performances between boys and girls in order to examine some of the problems with this reductionist approach.

The Fallacy of Biological Determinism[1]

Before I proceed to examine the scientific basis of the arguments about math ability, it is important to point out that even the best experiments which might show a genetic component to an ability such as math performance are of no use in devising educational policy. There is a widespread misconception that assigning a trait a genetic component means that the trait is fated to exist. But contrary to what many people believe, genetic does not mean determined—set in stone. Whether we were able to show that a particular trait was substantially biologically influenced or substantially environmentally influenced would tell us nothing about its changeability.

For instance, consider the *genetic* disease, phenylketonuria (PKU). Individuals suffering from PKU are unable to use the amino acid phenylalanine which is present in protein in the diet. Toxic derivatives of this amino acid accumulate in the body and brain, causing in most if not all cases very severe mental retardation. The disease is well-characterized and the specific genetic mutation which causes it is understood. However, once the basis of this disease was discovered, doctors realized that if they fed a newborn child a phenylalanine-free diet, the toxic compounds would no longer accumulate. Experience has shown that children treated with such a diet develop normally without any signs of the retardation previously observed.

Thus, there is a disease, the genetic origin of which is undeniable, but which is not by any means fated to persist. The difference between PKU individuals and individuals without this genetic mutation disappears when the environment is changed. This example illustrates a basic and very important concept in genetics. The statement that a particular property of an individual (in this case a disease) is inherited is only meaningful for the environment in which it is studied. The "genetically based difference" between individuals can vanish in a different environment. Similarly, people with an inherited defect in vision, in effect lose that defect when they wear glasses.

What relevance does this argument have for the debates over biological differences in intelligence or school performance? What it means is that if some studies should reveal a real genetic basis for some difference between girls and boys or blacks and whites (in either direction), they would not assist in the development of educational policy. Whatever those traits, the evidence that there is a biological difference only says that given current social, environmental, and educational techniques those differences exist. But a change in any of these environmental factors could eliminate the differences or even make the previously "inferior" group "superior."

For Benbow to say that women should accept their differences and for *Newsweek* to state "if they [the differences] are genetic, we must learn to accept them" (Williams & King, 1980) do not follow. Not only are claims for a genetic basis for some difference of no use in setting classroom priorities, they also tell us nothing about how we might go about eliminating those differences.

The Benbow and Stanley Study

The reader might think at this point that since the significance and usefulness of any studies which purport to show a genetic basis for differences in educational achievement are highly questionable, there is no point in going through the studies themselves to assess their validity. Whether they were poorly or well done, they would not be of much interest. However, I believe it is useful to examine such studies because they are good examples for understanding the social biases inherent in doing science. For, beyond the succumbing by authors such as Benbow and Stanley to the fallacy of biological determinism, the studies themselves seem inevitably to be fatally flawed by the strong preconceptions brought to them by their authors.

Let us examine this most recent study as an example of the problems inherent in attempts to detect genetic influences on human social behavior. In a research article entitled "Sex Differences in Mathematical Ability: Fact or Artifact?" Benbow and Stanley (1980) reported the results of eight years of research conducted by the Study of Mathematically Precocious Youth (SMPY). This group conducts talent searches to identify mathematically gifted children. Benbow and Stanley gave seventh and eighth grade children in each of several talent searches the Scholastic Aptitude Test (SAT) and found that on the mathematics section the boys, on the average, got higher scores than the girls. Since children of this age had presumably not learned in school the material covered by the SAT, Benbow and Stanley concluded that the test in this case measures true mathematical ability. In addition, since seventh and eighth graders have taken the same math courses, differences in achievement at this level cannot be attributed to boys having received more education.

Benbow and Stanley conclude:

> We favor the hypothesis that sex differences in achievement in and attitude toward mathematics results 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. (p. 1264)

Benbow and Stanley's conclusion is worded in such a way as to suggest to the reader that innate (or "endogenous" in their words) factors played an important role in the differences in math performance. The use of the phrase "superior male mathematical ability" certainly implies something inborn. Further, their brief allusion to "boy-versus-girl socialization processes" indicates an attitude that these "processes" are not the explanatory factors. This is certainly how the conclusion was taken by *Science* itself, which published an accompanying news feature entitled "Math and Sex: Are Girls Born with Less Ability?" (Kolata, 1980). Subsequent news articles in the mass media sported headlines such as "The Gender Factor in Math" (1980) and "Do Males Have a Math Gene?" (Williams & King, 1980).

The impression that Benbow and Stanley were proposing a genetic hypothesis to explain differences in math ability was reinforced by public statements of the researchers themselves. Benbow's comment that women would be better off accepting the differences and Stanley's suggestion that the observed differences might be due to different brain lateralization in boys and girls which could be genetically programmed (Brandt, 1981) indicate that the interpretation given by most observers to their circuitously worded conclusion was an accurate reflection of the researchers' attitudes.

The Assumptions behind the Study

Any scientific study must proceed with certain assumptions. Assumptions allow one to ignore hypotheses which seem extremely unlikely. In this way, when doing a scientific study, one does not have to control for an infinite number of unlikely factors which might be dreamed up to explain a phenomenon. As Ornstein put it:

> Science as a mode of knowing involves a limitation on inquiry. The essence of a good experiment is successful exclusion.... If for example, we want to study the response of cells in the brain to visual stimuli, we would be considered mad if we also monitor the blood flow to the feet, the temperature of the room, the phase of the moon, the growth rate of mushrooms outside, or any one of the millions of available possibilities. (Ornstein, 1975, p. 22)

This character of scientific investigation is as true of work in my own field, bacterial genetics, as it is in the study of human behavior. However,

there is a difference. The assumptions in the field of *human* behavior inevitably reflect the investigator's *social* perspective. As an example, the study of the "inheritance of IQ" requires, among other assumptions, the conviction that an IQ test is culturally unbiased and that it measures something real—a real property of individuals. These assumptions are not themselves supported by extensive scientific studies. They are acts of faith over which people with different social attitudes disagree. Thus, one of the primary ways of examining a study such as the one under consideration here is to determine what those assumptions are and how they influence the study and its conclusions. In the case of the Benbow-Stanley study, several critical assumptions are made. First, Benbow and Stanley appear to assume that differential socialization of girls and boys is not an important factor in determining the differences in math performance. The one causal factor other than innate math ability which these authors consider important enough to mention is differential course-taking. But this factor is not even considered seriously enough to present data. Their claims that the children had taken the same courses is referenced "C. Benbow and J. Stanley, manuscript in preparation." In this way, the authors ignore a whole body of research examining female/male socialization and its possible impact on math performance. This is particularly surprising when we consider that Stanley is general editor for a series of books which includes *Women and the Mathematical Mystique* (Fox and Cohn, 1980). This book reports on the work of a number of researchers who have documented social factors in school and at home affecting attitudes toward math.

Early childhood treatment and differential upbringing of the sexes could have significant effects on later interest and performance in math. The different kinds of toys that boys and girls are given to play with, boys' interest in sports which require the development of various mathematical reasoning powers, and parents' attitudes toward their children's school work could all have a significant impact. Casserly (1980, pp. 138–163) reported in a study of such parent-child interactions that "girls agreed that a chemistry set had been the hardest toy for most of them to get" (p. 152).

Fox and Cohn (1980, pp. 94–112) cite a study of gifted children which found that parents of gifted boys often noticed their sons' interest in science at an early age, discussed careers with them, and supplied them with science-related toys and books. Very few noticed their daughters' interest in science. Gessner (cited in Tobias, 1982) found that when parents were given an older SAT test and told that their children would be taking such a test in the near future, the parents were more likely to help prepare their sons than their daughters. Parsons (1982) concluded from a path analysis of math performance and various factors that mothers play a critical role in socializing children with regard to sex differences in mathematics attitudes and achievements.

The encouragement of boys and discouragement of girls in math has also been examined in the school setting. This discouragement process begins early. Ernest (1976), in a survey of elementary and high school teachers, found that 41% thought boys were better than girls at math, while none thought girls were better than boys. He suggested that "we may be observing the so-called Pygmalion effect in education, according to which the student performs to some (measurable) extent, in response to the expectations of the teacher" (p. 614).

Guidance counselors and teachers often discourage girls from taking math. In one study it was found that "42 percent of girls interested in careers in mathematics or science reported being discouraged by counselors from taking courses in advanced mathematics" (Haven, 1972, p. 5). Casserly (1980, pp. 138–163) interviewed guidance counselors and came up with comments along the lines of "I just hate to see a girl get in over her head." Even if the students do go on to further math courses, the message conveyed by these advisors is bound to affect their confidence in their mathematical ability.

At least two studies have directly examined actual differential treatment of girls and boys in the classroom. In 33 second grade classes, Leinhardt, Seewald, and Engel (1979) found that teachers made more academic contacts and spent more "cognitive time" with girls in reading and with boys in math. Becker (1981) showed that in high school geometry classes, teachers favored boys over girls in a number of ways including "cognitive levels of questions," "praise and criticism," "encouragement," and "individual help."

All of these studies speak directly to Benbow and Stanley's claim to have controlled for differential course-taking. Sitting in the same classroom and learning from the same teacher is clearly a different experience for girls than for boys.

Parental, teacher, and societal attitudes may also be responsible for the lessened confidence of girls in their math ability. Wolleat, Pedro, Becker, and Fennema (1980) found that girls were more likely than boys to attribute their success in math to hard work rather than ability. Conversely, girls would, more often than boys, invoke lack of math ability to explain their failures in math. This was true at every level of achievement. They also concluded from earlier work that girls as compared to boys are less confident about their math ability and consistently underestimate their ability to solve mathematical problems. Dornbusch (1974) found the same phenomena in a separate study.

Furthermore, girls may be socialized to not want to do well in math because boys might not like them or they might be socially ostracized. Interviews with girls have turned up the following representative comments:

> Boys do not like or are afraid of smart girls, especially math whizzes. (Luchins & Luchins, 1980, p. 13)
>
> Girls don't want to be known as science nuts by taking every science course around. (Casserly, 1980, p. 149)
>
> A teacher told me I would have a hard time finding a husband if I kept on doing so well in math.[2]

Some studies have suggested that teenagers associate mathematics with masculinity (Tomizuka & Tobias, 1981).

These many ways in which boys and girls are differentially socialized may also explain the greater involvement of boys in activities outside of the classroom which involve math skills (Burton, 1979). Extracurricular activity may, in fact, be a major factor leading to differential performance on the SAT tests. This is the conclusion drawn by Senk and Usiskin (in press), who studied female/male performance on geometry proofwriting with high school students. They found that while girls entered geometry courses knowing less geometry, by the end of the course there was no difference in performance between the sexes. This was true on the average as well as at the upper levels of performance. They suggest that the lack of difference may be explained by the fact that experiences in geometry are not usually of the type that students are likely to be involved in outside of school. They conclude that "the more an instrument directly measures students' formal educational experiences in mathematics, the less likelihood of sex differences."

One of the most striking findings of this and other studies (Armstrong, 1981; Stage & Karplus, 1981) is that on tests of spatial visualization there is no difference between girls and boys. It is clear that the supposed deficit of females in spatial visualization which seems to be taken as a given by Benbow, Stanley, and many others working on sex role differences is, to say the least, far from well-established.

Finally, an ironic contrast is provided by a look at a study carried out by Fox and Cohn (1980), two other investigators at SMPY. Using the same sample of gifted children, they found that, even in the seventh grade, the high-scoring boys have a strong orientation toward investigative careers in mathematics and the sciences, and a strong "theoretical-value orientation." The high-scoring girls tend to have values that are more social than theoretical. Furthermore, the boys much more than the girls (as in other studies) seek out extracurricular experiences in mathematics—studying with parent or teacher, working mathematical puzzles and so forth. Fox and Cohn conclude: "SMPY's study of the characteristics of mathematically precocious adolescents lends some support for the social explanation of sex differences at the higher levels of ability and achievement" (p. 100).

This conclusion is especially striking because the SAT performance data of Fox and Cohn are exactly the same as Benbow and Stanley's! Clearly, some sort of bias is causing a "differential" interpretation of data between the two groups.

Thus, there is an impressive body of literature examining a host of social factors which may affect math performance. I should not minimize the difficulty in studying these problems by claiming that these studies "demonstrate" that the differences between boys and girls are explained by such factors. But what this work does illustrate dramatically is the extraordinary complexity in the development of an ability such as math performance. It also illustrates the foolhardiness of attempting to attribute sex differences in math achievement to genetic factors. In fact, the only way to examine meaningfully the question of a genetic contribution in this case would be to equalize child-rearing practices, change teachers' attitudes, and alter other factors for which there is more than circumstantial evidence as influential parameters. Once the environments are equalized, if it is still of interest to examine these questions, then a considerably more meaningful setting for such studies will have been established.

Seen from this perspective, the most basic assumption behind these studies—that it is possible to carry out research which would find a genetic component of the differences—itself reflects a social bias. It requires that one rule out the entire boy-girl socialization process in and out of school as a significant component of the differences. In turn, the only basis for doing so is a preconceived notion that genetics plays a predominant role. This very important assumption predetermines the conclusions of the study.

Another assumption operating in the analysis of Benbow and Stanley may be a faith in those observations which can be quantified and analyzed statistically. Benbow and Stanley's reference to socialization arguments as the "usual armchair assumptions" (Benbow & Stanley, 1981, p. 147) in contrast to their extensive data on 20,000 students appears to reflect this faith. It is relatively easy to amass large numbers of test scores, to determine chi-squares, and so forth. It is not so easy to quantify social, parental, and teacher influences on confidence, expectations, and attitudes, and the effect of these influences on math performance. These problems do not invalidate arguments from such studies.

The SAT-M and Math Ability

A second major problem with the study by Benbow and Stanley is their reification of the SAT-M scores. According to paleontologist Stephen Jay Gould (Gould, 1981, p. 250), the error of reification in this case is to assign a physical meaning to a test score. That is, for SAT's, IQ, and other such tests, the assumption is made that the score is a measure of a unitary innate ability, in turn determined by brain cells and ultimately genes. But there is no basis for this assumption.

SAT's were developed and have evolved to be predictors of academic performance in college. Nowhere in the construction of these tests does a mechanism appear for connecting test scores with an innate ability. In fact, Jackson (1980), while defending the SAT against its critics, did admit that "the developers of the SAT do not view it as a measure of fixed capacities" (p. 383). This view has been forced on the producers of the SAT tests by the results of various studies on the effect of coaching on SAT scores. Various within-classroom and extra-classroom short-term coaching regimens have led to as much as an average 85 point increase in SAT-M scores (Slack & Porter, 1980a, 1980b). These changes in test scores point out that performance on a test is the result of a complex interaction between the brain and experience. While there had been no differential coaching of the boy and girl SMPY students, the host of socio-cultural factors I have described in the previous section provides at least as reasonable an explanation for test score differences as the genetic one.

Another assumption that must be made in order to consider that the SAT test is a measure of innate ability is that the tests are unbiased with regard to the sexes. However, performance in word problems is affected by the problem solver's familiarity with their content. For example, Graf and Riddell (1972) gave their subjects one of two word problems which were identical computationally. One was about buying fabric and the other about selling stocks. While girls and boys took equally long to solve the fabric problem, girls took much longer to solve the stock problem. Studies by the Educational Testing Service itself have found sex biases in the contents of the SAT word problems (Donlon, 1973).

The elimination of bias is an enormous problem in the design of tests. Consider the history of the Stanford-Binet IQ test. When it was first prepared in 1916, women had higher scores than men. Since the test makers apparently made a judgment that women and men were equal in intelligence, the tests had to be redesigned. Certain questions in which women did consistently better than men were eliminated so as to result in the same mean score for men and women (Salzman, 1977). The same people could also have decided, if they had wished, to readjust tests in order to have blacks and whites have equal scores. What is clear is that assumptions about the relative intelligence of groups of people have gone into the very design of the SAT and other tests.

Another factor which introduces a sex bias in test-taking performance is "math anxiety" (Tobias, 1978) which is more prevalent among girls than boys. (Interestingly, Benbow and Stanley [personal communication] report, in the follow-up on their SMPY students, that the girls who had done more poorly than boys on the SAT-M in their study did better in math courses in high school than the boys!)

Finally, there is an assumption in the Benbow and Stanley study and in their remarks to the press that scores on the tests are predictors of who will enter into careers based on mathematics. In fact, there is no correlation between SAT-M scores and future careers in math. Students in the SMPY study did not go on to graduate careers in mathematics, although many did enter graduate school in other subjects (Schafer & Gray, 1981). Parsons (1982) concludes from her analysis that attitudinal factors play a greater role than math aptitude in determining which students will take courses in mathematics.

The fate of the IQ controversy provides an interesting lesson for understanding the pitfalls of attempts to connect performance on tests with innate ability. Identical twin studies were used as a basis for drawing conclusions about inheritance of intelligence. These were extended to explain the differences in performance and achievement in society between blacks and whites and between socio-economic classes. Extensive scientific critiques of these studies which appeared shortly afterwards received little public attention. It was not until the extraordinary tale of fraud surrounding the career of Sir Cyril Burt was unraveled (Hearnshaw, 1979) that the media paid attention.[3]

Two studies done since the original publicity on IQ and genetics, which seem to undercut the entire foundation for the hereditarian point of view, have received very little attention. Scarr and Weinberg (1976) studied 130 black/interracial children adopted by white advantaged families. They found that, depending on the age of adoption, these children scored 16 to 20 points higher on IQ tests than black children in the general population. In France, Schiff, Duyme, Dumaret, and Tomkiewicz (1982) studied 32 children of unskilled workers who were adopted by families from the top 13% of the socio-professional scale. Compared to children of unskilled workers who remained with their natural families, including some siblings of the adoptees, the adopted children scored on the average 14 points higher on IQ tests.

These findings indicate the remarkable malleability of IQ test scores. Myriad factors may be responsible for the increase in IQ scores seen in these studies. For instance, the cultural biases built into such tests, proposed by many researchers, would provide a reasonable explantion for the results.

Neither of these studies has received significant media attention. Yet they would appear to be direct tests of predictions of the genetic hypothesis which have yielded results contrary to the theory. These results re-emphasize the many criticisms of the SAT and IQ tests, particularly the error of reification.

The Social Impact of the Benbow and Stanley Study

Julian Stanley, in an interview with *Science*, stated "We want our data out in the public domain so they can't be ignored" (Kolata, 1980, p. 102). The success in reaching "the public domain" has been impressive. In addition to major newspaper stories (Are boys better at math?, 1980; At mathematical thinking, boys outperform girls, 1980) and the *Time* and *Newsweek* articles referred to, *Discover* magazine began its cover story on "The Sexes and the Brain" (Weintraub, 1981) with reference to this same study as the most recent support for the position that sex role differences are inborn. According to the subtitle of this article, "Men and Women think differently. Science is finding out why." *Reader's Digest* devoted major space to the math study in an article on sex role differences (Durden-Smith & DeSimone, 1982). A section entitled "Math Gene" concludes that "such differences linked to brain organization may help to explain why members of one sex or the other are overrepresented in certain professions" (p. 266).

Given all the extraordinary limitations of the Johns Hopkins study, it is at first surprising to see the extent of the media's reaction and their willingness to broadcast speculations based on the flimsiest of evidence. In part, their interest may have been attracted by the fact that *Science* chose to feature the article with a provocatively titled news story.

However, the publicity this study received is consistent with other trends in science reporting on male-female differences during the last few years. Increasingly, we find scientists suggesting and the press reporting that women's current place in society—intellectually, economically, sexually, and in power relationships in general—is a natural consequence of differences in female and male biology.[4] Sociobiologists suggest that "mother nature is sexist" (Barash, 1977, p. 283) and that such social facts as the sexual double standard (Keen, 1981), heterosexual rape (Rhodes, 1981), and the dearth of women in "science, government and business" (Wilson, 1979, p. 138) are a consequence of natural selection which operated on males and females differently. Others propose that different brain structure or hormonal makeup between men and women leads to the different sex roles found in society, including differential math performance (Weintraub, 1981; Durden-Smith, 1980; Gelman, 1981).

It is, of course, not unreasonable to suggest that biology will have an influence on human male and female behavior. However, in all of these studies, as in the Benbow-Stanley study, there is simply no evidence to support the conclusions put forth. Sociobiology has been widely criticized as speculation built on the social biases of the scientists involved (Washburn, 1978; Lewontin, 1980; Kaplan, 1978; Leeds & Dusek, 1981–82). The evidence for significant differential brain lateralization and for any connection it might have with sex role differences is, at best, very weak (Springer & Deutsch, 1981). Even the most prominent researchers in the hormone work admit the difficulty of separating the influence of early childhood socialization from biological factors (Erhardt & Meyer-Bahlburg, 1981; Rubin, Reinisch, & Haskett, 1981). There have simply not been any new scientific breakthroughs or insights.

Rather, the publicity sex-role research has received—and possibly the research activity itself—can be explained as a social and political phenomenon. The problems to which biologists and other scientists have claimed to bring new insight correlate strikingly with the issues which have been raised by the women's movement in the last 10 to 15 years. Demands for equal rights and affirmative action are now met with the argument that biology limits women's possibilities. Increased consciousness of the problem of rape and its connection to power relationships between men and women are met with statements that rape and male dominance are natural consequences of the male's need to spread his genes as widely as possible. Seen in this light, science appears as a social weapon to be used to maintain current power relationships in the society.

I am not suggesting that there is a deliberate conspiracy to promote such research or to publicize it. Instead, the flourishing of such research may be a consequence of the influence of the dominant ideology on members of both the scientific community and the media.

What effects do such studies and the publicity surrounding them have on social attitudes and social policy? These are of course hard to measure. I have described earlier the impact of other biological-determinist arguments on educational policy. It also seems indisputable that reading headlines such as "Born Dumb?" (1969), after Jensen's article was published, had a serious influence on the confidence of blacks in their abilities and on the teachers' treatment of black students.

In the case of the Benbow-Stanley study, Fennema (1981) reports the results of interviews with sixth to eighth grade boys and girls after the study had appeared. Comments included the following:

> Boys are better. Cuz I've seen like studies and they say boys are better.
>
> Okay, cuz I read somewhere...that, um, boys are, it's some kind of scientific thing that boys are better in math than girls are. (p. 381)

I recently talked with a high school science class about the Benbow and Stanley study. One girl who had read stories about the study said that she felt as if she should no longer try to improve her grade in math. It also seems likely that such publicity can only worsen the differential treatment boys and girls receive.

It also is disheartening to see the publicity this study has received when there have been significant successes in the last 10 years or so with various programs designed to increase female participation and performance in math (MacDonald, 1980; Brody & Fox, 1980; Blum & Givant, 1980; Fennema, Pedro, Wolleat & Becker, 1981; Taylor, 1983). Taylor describes a program in the Minneapolis public schools begun in the mid-1970's. Workshops on sexism were held with teachers, visiting women scientists gave talks, teaching materials with sexist stereotypes were eliminated from the curricula, and various other programs were introduced to encourage girls to enter more advanced math courses. From 1976 to 1981 the percentage of girls taking calculus courses increased from 25 to 43%, while in math analysis it increased from 35 to 43%. In math courses overall the participation of girls was approaching 50%. While Taylor admits that this change may have been due to a combination of factors, including increased feminist consciousness, the change in girls' math interests is an encouraging finding whose significance cannot be vitiated by the Benbow-Stanley study. Furthermore, in the 10 years between 1970 to 1980, the number of women PhD's in mathematics increased from 6 to 15%. Within the field of study of women's math education, then, the Benbow-Stanley study would appear to fly in the face of everything else that is being learned. It is unfortunate that the reverse impression is given by the media.

Conclusion

Scientific research is not a neutral pursuit. We bring to our work our own social attitudes which inevitably play a role in the development of any research program. The questions we ask, the assumptions we make, the approaches we use, all are influenced by our social attitudes. I have presented here an analysis of how such biases enter into studies such as the one done by Benbow and Stanley. In the case of this particular study, those biases appear so strong as to render any conclusions from the study highly problematic.

Educators and the public are being confronted with more and more reports from the scientific community which are said to have implications for educational policy. It is important to be able to analyze these reports from a critical perspective. By looking first at the assumptions which underlie such studies, it is often possible to immediately detect the problems. Case studies such as this one can also be useful in high school or college classrooms to bring a more critical attitude toward scientific reports among students. Finally, it bears repeating that even if a study should come along which would overcome some of the apparently insuperable difficulties in studying the biological basis of complex human social behavior, the results in no way can be used to set social or educational policy. Biological or genetic does not mean fated.

Footnotes

[1]Lewontin (1976) gives an excellent summary of these arguments.

[2]This comment was made by one of the students in my graduate seminar class in genetics after a discussion of the Benbow and Stanley article.

[3]As described by Hearnshaw (1979) and others who have investigated the identical twin studies of Burt, this case represents one of the most extraordinary examples of fraud in the history of science, particularly considering the impact of Burt's work. It appears that Burt must have fudged his data and concocted for his publications the names of coworkers who have never existed. There is also a strong likelihood that the twins on which these studies were based never existed.

[4]Two recent books which evaluate critically arguments about biology and sex roles are Sayers, 1976, and Hubbard, Henifin, and Fried (1982).

References

Are boys better at math? *The New York Times*, December 7, 1980, p. 102.

Armstrong, J. Achievement and participation of women in mathematics: Results of two national surveys. *Journal of Research in Mathematics Education*, 1981, *12*, 356–372.

At mathematical thinking, boys outperform girls. *Washington Post*, December 5, 1980, p. A1.

Barash, D. P. *Sociobiology and behavior*. New York: Elsevier, 1977.

Becker, J. R. Differential treatment of females and males in mathematics classes. *Journal of Research in Mathematics Education*, 1981, *12*, 40–53.

Benbow, C., & Stanley, J. C. Sex differences in mathematical ability: Fact or artifact? *Science*, 1980, *210*, 1262–1264.

Benbow, C., & Stanley, J. C. Sex differences in math reasoning. *Science News*, 1981, *119*, 147.

Blum, L., & Givant, S. Increasing the participation of women in fields that use mathematics. *Math Monthly*, 1980, *87*, 785–793.

Born dumb? *Newsweek*, March 31, 1969, p. 84.

Brandt, R. On mathematically talented youth: A conversation with Julian Stanley. *Educational Leadership*, 1981, *39*, 101–106.

Brody, L., & Fox, L. H. An accelerative intervention program for mathematically gifted girls. In L. Fox, L. Brody, & D. Tobin (Eds.), *Women and the mathematical mystique*. Baltimore: Johns Hopkins University Press, 1980.

Burbank, R. P. Brain growth: Learning by spurts and starts. *The Massachusetts Teacher*, January/February 1981, pp. 13–16.

Burton, G. Regardless of sex. *The Math Teacher*, 1979, *72*, 261–270.

Casserly, P. L. Factors affecting female participation in advanced placement programs in mathematics, chemistry and physics. In L. Fox, L. Brody, & D. Tobins (Eds.), *Women and the mathematical mystique*. Baltimore: Johns Hopkins University Press, 1980.

Donlon, T. F. *Content factors in sex differences*. Research monograph #73-m. Princeton: Educational Testing Service, 1973.

Dornbusch, S. To try or not to try. *Stanford Magazine*. 1974, *2*, 50–54.

Durden-Smith, J. Male and female-why? *Quest/80*, October 1980, pp. 15–19, 93–98.

Durden-Smith, J., & DeSimone, D. Is there a superior sex? *Reader's Digest*, November 1982, pp. 263–270.

Erhardt, A. A., & Meyer-Bahlburg, H. F. L. Effects of prenatal sex hormones on gender-related behavior. *Science*, 1981, *211*, 1312–1318.

Ernest, J. Mathematics and sex. *American Mathematical Monthly*, 1976, *83*, 595–615.

Fennema, E. Women and mathematics, does research matter? *Journal of Research in Mathematics Education*, 1981, *12*, 380–385.

Fennema, E., Pedro, J. D., Wolleat, P. L., & Becker, A. D. Increasing women's participation in mathematics: An intervention study. *Journal of Research in Mathematics Education*, 1981, *12*, 3–14.

Fox, L. H., & Cohn, S. J. Sex differences in the development of precocious mathematical talent. In L. Fox, L. Brody, & D. Tobin (Eds.), *Women and the mathematical mystique*. Baltimore: Johns Hopkins University Press, 1980.

Gelman, D. Just how the sexes differ. *Newsweek*, May 18, 1981, pp. 72–83.

Gould, S. J. *Ontogeny and phylogeny*. Cambridge, Mass.: Harvard University Press, 1977.

Gould, S. J. *The mismeasure of man*. New York: W. W. Norton, 1981.

Graf, R. G., & Riddell, J. C. Sex differences in problem-solving as a function of problem content. *Journal of Educational Research*, 1972, *65*, 451–452.

Haven, E. W. Factors associated with the selection of advanced academic mathematical courses by girls in high school. *Research Bulletin*, 72–12. Princeton: Educational Testing Service, 1972.

Hearnshaw, L. S. *Cyril Burt, Psychologist*. Ithaca: Cornell University Press, 1979.

Hubbard, R., Henifin, M.S., & Fried, B. *Biological woman: The convenient myth*. Cambridge, Mass.: Schenkman, 1982.

Jackson, R. The S.A.T.: A response to Slack and Porter's "Critical Appraisal." *Harvard Educational Review*, 1980, *50*, 382–391.

Jensen, A. R. How much can we boost I.Q. and scholastic achievement? *Harvard Educational Review*, 1969, *33*, 1–123.

Kaplan, A. L. *The sociobiology debate*. New York: Harper and Row, 1978.

Keen, S. Eros and Alley Oop. *Psychology Today*, February 1981, pp. 52–61.

Kolata, G. B. Math and sex: Are girls born with less ability? *Science*, 1980, *210*, 1234–1235.

Leeds, A., & Dusek, V. Sociobiology: The debate evolves. A special issue of *The Philosophical Forum*, 1981–82, *13* (2–3).

Leinhardt, G., Seewald, A. M., & Engel, M. Learning what's taught. Sex differences in instruction. *Journal of Educational Psychology*, 1979, *71*, 432–439.

Lewontin, R. The fallacy of biological determinism. *The Sciences*, March/April 1976, pp. 12–15.

Lewontin, R. Sociobiology: Another biological determinism. *International Journal of Health Services*, 1980, *10*, 347–364.

Loviglio, L. Mathematics and the brain: A tale of two hemispheres. *The Massachusetts Teacher*, January/February 1981, pp. 9–12.

Luchins, E. H., & Luchins, A. S. Female mathematicians: A contemporary appraisal. In L. Fox, L. Brody, & D. Tobin (Eds.), *Women and the mathematical mystique*. Baltimore: Johns Hopkins University Press, 1980.

MacDonald, C. T. An experiment in mathematics education at the college level. In L. Fox, L. Brody, & D. Tobin (Eds.), *Women and the mathematical mystique.* Baltimore: Johns Hopkins University Press, 1980.
Ornstein, R. *The Psychology of consciousness.* New York: Penguin, 1975.
Parsons, J. E. Social forces shape math and attitudes and performance. *Association of Women in Mathematics Newsletter,* 1982, *12* (3) 4–9.
Rhodes, R. Why do men rape? *Playboy,* April 1981, p. 112.
Rubin, R. T., Reinisch, J. M., & Haskett, R. F. Postnatal gonadal steroid effects on human behavior. *Science,* 1981, *211,* 1318–1324.
Salzman, F. Are sex roles biologically determined? *Science for the People,* July/August 1977, pp. 27–32.
Sayers, J. *Biological politics: Feminist and anti-feminist perspectives.* New York: Tavistock, 1982.
Scarr, S., & Weinberg, R. I.Q. test performance of black children adopted by white families. *American Psychologist,* 1976, *31,* 726–739.
Schafer, A. T., & Gray, M. W. Sex and mathematics. *Science,* 1981, *211,* i.
Schiff, M., Duyme, M., Dumaret, A., & Tomkiewicz, S. How much *could* we boost scholastic achievement and I.Q. scores? A direct answer from a French adoption study. *Cognition,* 1982, *12,* 165–196.
Senk, S., & Usiskin, Z. Geometry proofwriting: A new view of sex differences in mathematics ability. *American Journal of Education,* in press.
Slack, W. V., & Porter D. R. The Scholastic Aptitude Test. A critical appraisal. *Harvard Educational Review,* 1980, *50,* 154–175. (a)
Slack, W. V., & Porter, D. R. Training, validity and the issue of aptitude: A reply to Jackson *Harvard Educational Review,* 1980, *50,* 392–401. (b)
Springer, S. P., & Deutsch, G. *Left brain, right brain.* San Francisco: W. H. Freeman, 1981.
Stage, E. K., & Karplus, R. Mathematical ability: Is sex a factor? *Science,* 1981, *212,* 114.
Taylor, B. R. Equity in mathematics. A case study. *Mathematics Teacher,* 1983, *76,* 12–17.
The gender factor in math. *Time,* December 15, 1980, p. 57.
Tiger, L. Male dominance? Yes, alas. A sexist plot? No. *New York Times Magazine,* October 25, 1970, pp. 35–37, 124–127, 132–136.
Tobias, S. *Overcoming math anxiety.* New York: Norton, 1978.
Tobias, S. Sexist equations. *Psychology Today,* January 1982, pp. 1614–1617.
Tomizuka, C., & Tobias, S. Mathematical ability: Is sex a factor? *Science,* 1981, *212,* 114.
Washburn, S. Human behavior of other animals. *American Psychologist,* 1978, *33,* 405–418.
Weintraub, P. The brain: His or hers. *Discover,* April 1981, pp. 15–20.
Williams, D. A., & King, P. Do males have a math gene? *Newsweek,* December 15, 1980, p. 73.
Wilson, E. O. *On human nature,* New York: Bantam, 1979.
Wolleat, P. L., Pedro, J. D., Becker, A. D., & Fennema, E. Sex differences in high school students. Causal attributions of performance in mathematics. *Journal of Research in Mathematics Education,* 1980, *11,* 356–365.