velop increasingly large gaps in their success rates for familiar and unfamiliar types of problems, relative to the gaps shown by less cautious problem solvers. When faced with the repeated and varied challenges of a test such as the SAT, they may avoid approaches such as working with unfamiliar modes of representation that seem risky but that offer the best chance of success. The influence of this sort of developmental process on mathematical reasoning needs to be understood before sex differences in performance are attributed by default to biological factors.

When sex differences in mathematics performance are understood in terms of both problem-solving and developmental processes, it may be possible to provide the "remedial" instruction for girls that Benbow suggests. However, a process model of mathematical reasoning may provide evidence that girls' reasoning processes are different from, but not necessarily inferior to, those of boys. The current higher mathematics has been constructed almost entirely by men who were doing the best thinking of which their male minds were capable (Keller 1985). This mathematics is beautiful and useful, but it is not the only mathematics that could exist. Perhaps, as more women go far enough within the male system to do creative work of their own, they will produce new conceptualizations and operations that take advantage of characteristically female modes of reasoning that have yet to be identified.

Sex differences in arithmetic computation and reasoning in prepubertal boys and girls

Arthur R. Jensen
School of Education, University of California, Berkeley, Calif. 94720

In her review of the evidence for sex differences in mathematical reasoning ability among intellectually gifted students, Benbow states: "We do not know how these findings may relate to students of average ability. Differences may be smaller at that level." She also states: "It is in junior high school that the sex difference in mathematics first becomes apparent. Girls excel in computation, boys on tasks requiring mathematical reasoning." She notes that, although the sex difference in mathematical reasoning is apparent by age 12, "It is rather difficult to obtain data below that age since there are no tests of mathematical reasoning ability for younger students, probably because the elementary curricula tend to cover mainly computation and basic arithmetic facts."

I have obtained data, not previously published, that throw some light on these and other issues posed in Benbow's target article. They consist of scores on three subtests (Arithmetic Computation, Arithmetic Concepts, and Arithmetic Applications) of the Stanford Achievement Tests battery obtained on *all* of the 3,112 pupils attending regular classes in the fourth, fifth, and sixth grades of all the elementary schools in one California school district. Virtually all of the boys and the vast majority of the girls are prepubescent, averaging about 9, 10, and 11 years of age in the three grades. The Computation test requires a knowledge of numeric facts and the various arithmetic operations known as mechanical arithmetic. The Concepts and Applications tests involve knowledge of quantitative concepts and the use of such concepts in reasoning about quantitative "thought problems" in which the required arithmetic operations are not explicit but must be inferred.

Table 1 shows the mean standardized sex differences (male minus female) on these tests. The reported difference (D) is what Benbow refers to as "effect size," that is, the raw-score mean difference divided by the average of the within-groups standard deviations. Also shown is the variance ratio (F) for male/female. In all three ethnic groups, girls perform very significantly better than boys on Computation (with an average effect size of −.20). On the Concepts and Applications tests, however, white boys significantly outperform their female counterparts (average effect size of +.185). The sex difference in the Asian (Chinese and Japanese) group favors girls in grades 4 and 5, then reverses in grade 6. The number of Asians is small, however, so we cannot make too much of this finding. The black pupils (total $N = 1{,}282$) show hardly any sex difference (average effect size of +.01) on Concepts and Applications, which agrees with Benbow's statement that the sex difference in mathematical reasoning ability is smallest among blacks. Also, only in the white group is the male/female variance ratio (F) consistently greater than 1. It is never significantly greater than 1 (overall mean $F = .95$) in the black group, even with its large N. The marked ethnic difference in the magnitude of the sex difference raises the question of whether this effect is attributable to cultural or biological factors. Benbow seems to favor the hypothesis that the sex difference in mathematical reasoning ability in her white sample is attributable to a biological difference between the sexes. But how would she explain the

Table 1 (Jensen). *Standardized mean male/female difference (D) and male/female variance ratio (F) on Arithmetic Computation, Concepts, and Applications of the Stanford Achievement Test taken by 3,112 elementary school pupils*

		White			Black			Asian		
		Grade			Grade			Grade		
Test		4	5	6	4	5	6	4	5	6
Computation	D	−.11*	−.08	−.27**	−.13**	−.29**	−.23**	−.41**	−.42**	.11
	F	.95	1.17	1.15	.93	.61	1.23	1.66	.69	1.00
Concepts	D	.17**	.21**	.21**	.09	.00	.02	−.33*	−.15	.16
	F	1.09	1.29*	1.28*	1.08	.63	1.12	1.73*	.63	.74
Applications	D	.18**	.18**	.16**	.00	−.10*	.10*	.32**	−.15	.28*
	F	1.12	1.43**	1.19	1.13	.69	1.03	1.94*	.85	1.00
N (Males)		269	274	280	218	219	226	36	44	40
N (Females)		264	223	278	212	216	191	43	43	36

Note: D = the difference between male and female means (M − F) divided by the average within-group standard deviation; F = the variance of males divided by the variance of females.
*$p < .05$, 2-tailed test.
**$p < .01$, 2-tailed test.

absence of a sex difference in mathematical reasoning in the present black sample? If the absence of a sex difference in the black group or the interaction of sex difference with ethnic group is explained strictly in terms of cultural factors, perhaps cultural factors also account for the sex difference in the white population. The observed interaction of the sex difference with ethnic group would seem quite problematic for Benbow's biological theorizing.

The male superiority in Arithmetic Concepts and Applications cannot be attributed to a sex difference in general intelligence. The Lorge–Thorndike Intelligence Test was given to all of these pupils, and in every ethnic group, at every grade level, the mean IQ of girls is about 3 IQ points *higher* than the mean IQ of boys – on both Verbal IQ *and* Nonverbal IQ. The overall male–female variance ratios are $F = 1.27$ for Verbal IQ and $F = 1.18$ for Nonverbal IQ (both F's significant at $p < .001$).

Most of Benbow's statistics are based on students who are highly selected for mathematical talent. What do we find when we select the top talent from the white group in the present study? If we select from above a cutoff 2 standrad deviations (*SDs*) above the overall mean of the total white distribution of combined scores on the Arithmetic Concepts and Applications tests, the ratio of the proportion of boys to the proportion of girls falling above the cutoff is approximately 4 to 1 (4% of boys and 1% of girls). If we select all pupils who score above a cutoff that is 2 *SDs above the mean for the boys' and the girls' distributions separately, so that approximately the same proportions of boys and girls are selected, the boys' mean turns out to be about two-thirds of a SD* higher than the girls' mean, which fully accords with Benbow's report of the effect size for the sex differences on tests of mathematical reasoning ability in groups of adolescents who are highly selected for mathematical talent.

Benbow's apparent hope for environmental remediation of the sex difference in mathematical aptitude seems to me a farfetched fantasy, if we believe that the hypothesis of a biological basis for the sex difference is correct. Remediation of a biological nature would raise the ethical question of the desirability of manipulating hormonal or other biological factors to achieve gender equality in high-level mathematical talent. Only an extremely small fraction of the male population displays the very high level of mathematical talent at which sex differences are marked. Therefore I would question the desirability, either to women or to society in general, of attempting to remedy a gender gap of such small consequence if the achievement of gender equality involved applying extreme or strenuous educational or biological interventions to any large segment of the population.

Biology: Si! Hard-wired ability: Maybe no

Douglas T. Kenrick

Department of Psychology, Arizona State University, Tempe, Ariz. 85287

It is interesting that Benbow opens her review by commenting on the problems of an ideologically based psychology of sex differences. Although times are changing, there are still those who believe that feminist ideology (or a particular sort of feminist ideology) should take precedence over data on gender differences. Advocates of this ideological approach have often used a two-front denial strategy. The first line is to deny that there are any substantial gender differences in behavior. This failing, the second line is to assume, a priori, that any differences that are demonstrated do not have a biological basis (e.g., Hubbard et al. 1982). Benbow does a fairly thorough job of vanquishing the denial of sex differences in math at the first front – it is hard to defend the position that there is no real gender difference here. She also makes a good attack against the second line of denial – the biological data are, if not overpowering, at least stronger than the environmental data.

Whatever one's feelings about a feminist-based approach to psychology, however, there seems cause for concern at the implications of this evidence – that mathematical problem-solving ability is biologically based. Would it not be better to suppress these data than to let the good ol' boys down at the legislature get the impression that women are innately inferior to males? I think not. There are two problems with the denial strategy, however instrumental it may seem at first. For one, it requires denial of other parts of the biologically based literature on sex differences, a literature that, in total, is hardly flattering to the good ol' boys. Along with any effects that male physiology may have on math and spatial ability, it seems to carry with it a wealth of handicaps. These include the learning disabilities and immune deficiencies discussed by Benbow, as well as a testosterone-fed proneness to violence (Rose et al. 1971) that, among other problems, leads males to be several times more likely to murder or to be murdered (U.S. Department of Justice, 1979; Wilson & Daly 1985). The other problem is that if we deny gender differences, we will not make any progress toward understanding, and thus toward remedying, those differences that displease us.

Benbow implies that the difference in math performance is due to a testosterone-induced modification in brain architecture. However, a look at the total picture emerging from biosocial research into sex differences suggests that the difference in math performance may not be due to a cognitive advantage for males at all. It may be a byproduct of other, much simpler gender differences in motivation and social behavior. Consider the clearest gender differences found to date – males are more aggressive, more oriented toward social-dominance games, and more physically active (see Eaton & Enns 1986; Kenrick 1987; for recent reviews relevant to these issues). Anyone who has seen Irven DeVore's fascinating film *The Baboon Troop* has watched the same pattern in our baboon cousins (see also Hall & DeVore 1965), and other primate research demonstrates the importance of testosterone in this masculine syndrome (e.g., Rose et al. 1971). Perhaps, then, the superior male performance on the SAT-M, and later in math-related professions, is simply another manifestation of the primate male's tendency toward hyperactive competitiveness.

Such an explanation fits with the finding that females do better than males under the more relaxed schedule of the formal math class, but worse under the time pressure of the SAT-M. In my own recollection, the quantitative sections of tests like the SAT were not so much more difficult than the verbal sections in content as they were more time pressured – and thus more likely to elicit a frantic race to finish before the bell. Even in grammar school, the boys would race to be the first to solve math problems, whereas the girls were more likely to quietly look on (even when they knew the answer). Performance on timed math exams may be like speed chess (a game that I have never seen a woman play, but which provides a peculiar obsession for several male colleagues). Thus, it may be that the SAT-M differences are simply a function of boys' higher motivation to perform on a timed competitive test.

Those higher levels of motivation could also account for the fact that the same boys go on to high levels of achievement in math and sciences. It seems unnecessary to presume that males' professional achievements are due to inherent abilities in those areas, since males' hyperactive dominance drives show up even in areas where there is no male ability advantage. In fact, males enter more competitively into the dominance hierarchies not only in physics and math, but also in art and literature, areas where female aptitude is at least on a par with that of males. In the math area, the dominance hierarchies are simply made relevant earlier because the tests draw out the male's affinity for "races." Consistent with my speculation here, Eccles (1985) reports reliably more persistent and single-minded pursuit of