On the Rarity of Mathematically and Mechanically Gifted Females:

A Life History Analysis

A dissertation submitted
by

Patricia Hausman

to

THE FIELDING INSTITUTE

in partial fulfillment of
the requirement for the
degree of

DOCTOR OF PHILOSOPHY
in
HUMAN DEVELOPMENT

This dissertation has been
accepted for the faculty of
The Fielding Institute

Abstract

On the Rarity of Mathematically and Mechanically Gifted Females: A Life History Analysis

by

Patricia Hausman

Engineering and certain physical sciences demand high levels of both mathematical and mechanical (HMHM) ability--a cognitive pattern found primarily among males. A small number of females also demonstrate this pattern. However, its correlates have not been examined longitudinally.

This study compared life histories of females with the HMHM pattern to those of other college-capable women. Using a model adapted from Helmuth Nyborg's theory of general trait co-variance, it predicted that HMHM females would have characteristics suggesting low lifetime exposure to estrogens--or to a high androgen/estrogen ratio. Subjects were 127 females from the National Longitudinal Survey of Youth.

Most somatic and reproductive predictions were supported. HMHM females matured more slowly than controls and were taller, thinner, and more physically active. Reproductive histories proved particularly noteworthy. HMHM females lost almost 25% of their pregnancies to miscarriage or stillbirth, and almost half were childless as of their early to late 30s. Controls had more pregnancies and births--and much lower rates of childlessness and pregnancy loss. Limited data on contraceptive use did not explain the fertility differential. HMHM females showed less religiosity than controls, but other psychological predictions were inconclusive or not supported.

Follow-up analysis considered whether study variables co-varied with general ability. Both groups were compared to a third (HIIQ) group equal to HMHM females in general ability but lacking marked mechanical aptitude. Means for HIIQ females on somatic and reproductive traits were generally intermediate to those of HMHM and control groups. By contrast, HMHM males showed some reproductive advantage over HIIQ males.

The results indicate that HMHM females differ biologically from controls and are consistent with reports that sex hormones influence cognitive architecture. The findings further suggest that the rarity of the HMHM pattern in females is best explained by the Darwinian principle of sexual selection. Evolutionary pressures select against characteristics that inhibit reproductive success. In females, factors associated with the HMHM pattern appear to fall into this category.

The limitations of the study, recommendations for further research, and suggested modifications to the study model are discussed. The need to replicate the findings in larger populations is stressed.

for my husband, Glenn

Acknowledgments

This study is based on data from the National Longitudinal Survey of Youth (NLSY). I would like to thank all of the subjects who participated in this extraordinary endeavor, as well as the researchers and interviewers who designed the survey and collected the data. I especially want to thank Steve McClaskie of the Center for Human Resource Research, who answered my many queries about the NLSY. His encyclopedic knowledge enabled me to negotiate its vast storehouse of information.

I am honored that David Geary of the University of Missouri agreed to serve as the external examiner for this work. He gave generously of his time and vast expertise in cognitive and evolutionary science, for which I am very grateful.

Marie Farrell, who chaired my dissertation committee, devoted many hours to this project and has my heartfelt thanks. I want to express appreciation to the other faculty who worked with me as well: Miguel Guilarte, Henry Braun, and Charles Seashore.

Thanks are also due to Leslie Wheeler, my student reader; Romana Mancini, who assisted with library research; Elyse Kutz, who oversaw administrative details from start to finish; and Warren Sarle, who was always willing to provide expert statistical advice.

This research was inspired by the work of Helmuth Nyborg, developer of the theory of general trait co-variance. Although he had no knowledge of this project, I would like to acknowledge the critical role that his outstanding contributions to psychology have played in my thinking.

Most of all, I want to thank my husband, Glenn Marcus, for his unflagging support during my years of doctoral study--as well his own expert input regarding countless technical aspects of this research.

Table of Contents

Figures

Tables

Chapter 1: Introduction

The tendency for humans to vary in their expression of countless traits is the raison d'etre for the science of differential psychology (Vernon, 1994). Among the attributes that vary among individuals are personality factors (Cloninger, 1987; Eysenck, 1981; Svrakic, Przybeck, & Cloninger, 1991), response to stimuli (Bettencourt & Miller, 1996; Halperin et al., 1995; Heath, Bresolin, & Rinaldi, 1989; Rowe, 1994; Steen, 1996), and achievement of developmental milestones (Andersson-Ellstrom, Forssman, & Milsom, 1996; Cooper, Kuh, Egger, Wadsworth, & Barker, 1996; Georgiadis, Mantzoros, Evagelopoulou, & Spentzos, 1997; Vanderzanden, 1989). Researchers have long sought to understand the origins of variation in characteristics such as these.

Nonetheless, cognitive ability appears to be the trait that has most fascinated researchers in the field (Gardner, Kornhaber, & Wake, 1996). In both magnitude and pattern, it varies substantially among individuals. Owing to researchers' longstanding interest in this phenomenon, cognitive ability may well be the best-studied aspect of human development (Halpern, 1992; Jensen, 1998; Kline, 1991).

One of the more enduring legacies of these inquiries is that cognition is influenced by an individual's sex. Although many ability tests are designed with the goal of minimizing sex differences in outcome (Jensen, 1980; Stumpf, 1995; Nyborg, 1994), domain-specific differences nonetheless occur with some frequency.

The number of studies reporting sex differences in particular abilities changes almost daily. Nonetheless, a reasonable estimate would be well into the hundreds--with the total number of individuals studied exceeding several million. Commentary prompted by the results of these investigations has created a burgeoning literature. Halpern (1992, p. 20) estimated that thousands, if not tens of thousands of journal articles and books address some aspect of this issue.

Typically, these findings show superior mean performance by females on certain measures of verbal ability--particularly tasks that require word fluency, reading comprehension, and perceptual speed (Halpern, 1997). For various mathematical competencies, an opposite and more exaggerated pattern that favors males has emerged (Felson & Trudeau, 1991; Jensen, 1998; Maccoby & Jacklin, 1974; Stumpf, 1995; Wilder, 1996).

Male superiority is particularly striking on measures of mechanical comprehension and spatial ability (Nyborg, 1983; Silverman, Phillips, & Silverman, 1996; Hedges & Nowell, 1995; Stanley, 1993; Geary, Gilger, Elliott-Miller, 1992). Many have long considered these aptitudes to be related to mathematical ability (Burnett, Lane, & Dratt, 1979; McCallum, Smith, & Eliot, 1979; Stanley, 1993). Geary (1996) has proposed an integrated model of this relationship. It provides a conceptual basis for appreciating how sex differences in spatial ability act through other biological and psychosocial factors to affect mathematical ability.

Compelling evidence of this relationship comes from the work of Casey and colleagues. These investigators showed that controlling for spatial skill among high-ability test-takers eliminated sex differences on the mathematics section of the Scholastic Aptitude Test (Casey, Nuttall, Pezaris, & Benbow, 1995). Others have also found spatial ability to be associated with better performance on this section of the test (Benbow, 1988). Mental rotation ability--the spatial skill that differs most among the sexes--predicts scores for females of both high and low academic ability (Casey et al., 1995).

The pattern of sex differences described above exists across a wide range of cultures (Kelly, 1978; Kempner, 1990; Silverman et al., 1996). However, the magnitude of difference in mathematical ability shows some variation. According to one report, it is not as great in Asian populations as in the United States (Lubinski & Benbow, 1992). Nonetheless, a reversed pattern in which men show less quantitative or spatial ability than women has yet to be reported.

A particularly interesting aspect of these findings has been the shape of the ensuing dialogue. Although some complain that educators, policymakers, and society ignore the needs of females (AAUW, 1991), findings of sex-related differences in cognitive performance have not fueled massive concern about the lower mean ability of males in certain verbal aptitudes. By contrast, the perceived mathematical deficits of females have prompted extensive commentary.

Ironically, most of this discussion has been limited to two statistical concepts. The first of these is the measure of central tendency described above. The second is a focus--perhaps uniquely human--on the so-called tail region of the ability curve. More specifically, this concern is directed at the rightmost end of the distribution--the area representing the highest levels of ability.

The numbers relevant to these concepts deserve mention. The difference in mean mathematical ability as measured in a cohort representative of the general population favors males by one-fifth to two-thirds of a standard deviation (Jensen, 1980). This magnitude of male advantage varies, of course, with the particular test--as well as with the age of the test-takers (Halpern, 1992; Willingham & Cole, 1997). The figures cited here are based on results from tests taken around the time of graduation from high school.

The magnitude of male-female difference is more dramatic at the highest end of the ability distribution (Feingold, 1988). At the 90th percentile or higher, males outnumber females by almost 2:1 (Hedges & Nowell, 1995). This ratio becomes progressively more unbalanced at increasing extremes of the ability spectrum. In the top percentile of mathematical reasoning ability, males reportedly outnumber females by 13:1 (Benbow & Lubinski, 1993). This estimate may be somewhat biased, however, because subjects were drawn from volunteers who participated in talent searches.

The distribution is significantly more skewed for spatial ability and mechanical comprehension. For these, the average values for males and females generally differ by a full standard deviation, or almost so (Hedges & Nowell, 1995; Jensen, 1980; Silverman et al., 1996; Stanley, 1993). As a consequence, males are dramatically overrepresented among those scoring in the highest decile on tests of these abilities. For example, on a standardized test of mechanical comprehension, males are statistically overrepresented by a factor of eight (Hedges & Nowell, 1995).

Meaningful as these data may be, the focus on a limited range of statistical measures has obscured important facts about the relationship between sex and mathematical ability. One of the more important of these is that males are statistically overrepresented not only at the highest levels of ability, but also at the lowest (Lehrke, 1997; Willingham & Cole, 1997). Simply put, females are less likely than males to be blessed with extraordinary talent or plagued by substantive deficits in mathematical ability.

Moreover, the magnitude of variation in quantitative ability is far greater within the sexes than between them (Wilson, 1992). As a result, a sizeable percentage of females achieve higher scores than the average male. This fact is acknowledged all too rarely. As sex differences are more pronounced for spatial and mechanical ability than for general quantitative ability, the number of females who outperform the average male on these aptitudes is much smaller. Nonetheless, some do.

Understanding the factors associated with a high degree of female competence in these domains is an important piece of unfinished business in differential psychology. Accordingly, this dissertation sought to identify characteristics associated with high mathematical and mechanical ability in females. It was also hoped that these correlates might provide insight into why so few women are gifted in both domains.

Significance

Research in cognitive psychology typically focuses on general intelligence or a single subtype of ability. By examining a particular combination of abilities, this investigation took a decidedly different approach. The reason for doing so was simple. Success in certain professional fields requires not simply a high level of verbal or mathematical skill, but, rather, a constellation of abilities.

This pattern studied here--marked by high ability in both mathematical and mechanical domains--predicts likelihood of success in fields such as engineering (DMDC, 1998; Hedges & Nowell, 1995; Stanley, 1993) and the physical sciences (Humphreys, Lubinski, & Yao, 1993).

Females are statistically underrepresented in all engineering fields. In 1993, they accounted for 15% of U.S. graduate students pursuing some type of engineering degree (NSF, 1994). In the subdisciplines of aerospace and mechanical engineering, however, they accounted for fewer than 10% of graduate students. A 1985 analysis of sex segregation in Danish universities similarly found females to account for only 10% of engineering students (Nyborg, 1994). This outcome followed years of well-orchestrated efforts by private and public agencies to increase female participation in male-dominated fields.

Moreover, enrollment figures can exaggerate actual outcome. Earning a degree does not guarantee employment in one's field of study. Actual employment of females as engineers is about half the reported rate of enrollment in graduate programs. In 1995, for example, females accounted for only 8% of individuals employed as engineers (Wooten, 1997).

The percentage of females in other physical sciences is somewhat higher, but far below those of males. Data for 1993 show that only about one-fourth of graduate students in the physical sciences were female. Their numbers were lowest in physics, where they accounted for only 14% of students.

Differences in interests partly explain the underrepresentation of females in these fields (Lubinski & Humphreys, 1990; Willingham & Cole, 1997). Females generally find people (and other living things) more interesting than inanimate objects (Lippa, 1998b). Nonetheless, the low number of females with the requisite combination of both mathematical and mechanical talent clearly remains an important factor (Hedges & Nowell, 1995; Humphreys et al., 1993).

The choice to combine mathematical and mechanical ability can be criticized on grounds that the latter does not correlate nearly as much with general intelligence as does mathematical skill (Jensen, 1998; Ree & Carretta, 1994). This point is well taken. Nonetheless, any drawback associated with studying this combination of cognitive traits must be weighed against its value.

This investigation revealed correlates of this cognitive pattern in females that have heretofore been unreported. In addition, the findings provided new insights into the origins of cognitive differences that contribute to sex segregation in the physical sciences and engineering.

Choosing a Theoretical Model

Research regarding within-sex differences in mathematics-related abilities among females comprises but a minute fraction of the literature on cognitive ability. Thus, no well-developed theoretical model for explaining within-sex variation was available. Fortunately, the large body of data on between-sex differences provided a basis for choosing among competing social, psychological, and biological models.

One of the more salient facts gleaned from this literature is that sex-related disparities in certain cognitive abilities are not present at every stage of life, nor of constant magnitude. Rather, differences that favor males are linked, in part, to the onset of puberty (Jensen, 1980).

The literature documenting this phenomenon is compelling (Felson & Trudeau, 1991; Wilder, 1996). In earlier grades, overall mathematical ability varies little between girls and boys. In fact, young females often outperform their male counterparts on standardized measures of mathematical ability (Jensen, 1998; Maccoby & Jacklin, 1974; NAEP, 1996). In light of these findings, conventional wisdom has long held that no math gap exists prior to the pubertal years.

However, two reports suggest that boys may have an advantage in specific areas of mathematics from the early elementary grades (Lummis & Stevenson, 1990; Mills, Ablard, & Stumpf, 1993). If confirmed, the notion that male superiority in mathematics does not emerge until puberty or the years just prior to it will no doubt be revised.

However, such a change would modify rather than reverse current thinking. While some data suggest that boys outperform girls in certain areas from the early grades, they also show that the magnitude of this advantage increases with the onset of puberty (Lummis & Stevenson, 1990). Accordingly, these findings do not contradict earlier notions linking sex differences in ability to sexual maturation.

In the related domain of spatial ability, studies of young children are somewhat inconsistent. Several reviewers have concluded that young boys outperform girls in some studies, but not others (Kerns & Berenbaum, 1991; Maccoby & Jacklin, 1974). Differences among various test instruments may account for some of the conflicting data. For example, the sex difference favoring boys is more evident on tasks requiring complex spatial reasoning than on simpler tests (Matthews, 1987).

Investigators who employed a comprehensive battery of seven different spatial tests concluded that the female disadvantage is evident around the age of 10 (Johnson & Meade, 1987). This timing coincides with the rise of estrogenic hormones in girls (Sizonenko, 1989). As is the case with mathematics, the full magnitude of sex differences in spatial ability is evident around the time of puberty (Jensen, 1980; Kerns & Berenbaum, 1991).

The magnitude of sex differences in spatial ability is greater in some areas than others. As noted earlier, the most dramatic difference occurs on tasks requiring mental rotation--the mental manipulation of objects in three-dimensional space (Geary, 1995). Most other types of spatial tests produce smaller, yet significant degrees of female disadvantage (Jensen, 1998; Kempner, 1990). However, females outperform males in one domain--the ability to recall the location of objects in a spatial array (Eals & Silverman, 1994; Silverman & Eals, 1992). For simplicity's sake, this so-called memory for location is excluded from the definition of spatial ability used here.

Unfortunately, data on sex differences in mechanical ability are not available for students in the elementary grades. These aptitude tests are rarely administered in the educational setting (Stanley, 1993). However, investigators have confirmed that the large sex difference in mechanical reasoning ability found among young adults also exists among gifted 12- and 13-year-old students (Stanley, 1993).

Both quantitative and spatial abilities contribute to mechanical aptitude (Bock & Moore, 1986; Kass, Mitchell, Grafton, & Wing, 1982). The mechanical ability factor measured on certain aptitude tests is sometimes referred to as Cox's *m*, and some believe it is comprised largely of spatial ability (Vernon, 1950). One might therefore expect sex differences in mechanical ability to increase around the time of puberty. However, absent data to evaluate this point, such a pattern cannot be assumed.

The nature of the adolescent-onset sex difference is also important. Many assume that the disparity results from girls failing to acquire new and more advanced skills at the same rate as males. This may not be the case. Longitudinal analysis of mathematics performance in elementary school has documented that girls lose their early advantage because familiar tasks become more difficult (Marshall & Smith, 1987). Specifically, girls had more difficulty in the sixth grade with tasks involving counting, visualization, and word problems than they did in third grade.

In short, the pattern of female cognitive ability appears to undergo a modest, but distinctive shift during the course of development. In areas where they performed as well or better than boys in early grades, females begin to underperform around the time of puberty. In domains marked by longstanding male advantage, the magnitude of female underperformance increases. This temporal factor suggests that an appropriate theoretical model is one that envisions sex hormones as a pivotal influence on the abilities in question.

The Theoretical Model

The model for this research was adapted from the theory of general trait covariance-androgen/estrogen, or GTC-A/E (Nyborg, 1994). Hereinafter, this model of human development is referred to as GTC theory, or Nyborg's theory.

GTC theory challenges the widely held view that social factors determine sex-typical traits such as occupational interests and personality. It posits that all gender-related characteristics--whether mental or somatic--develop harmoniously as a primary function of circulating sex hormones (Nyborg, 1984, p. 493). These circulating hormones include those secreted during gestation, which may have permanent effects on cerebral organization.

According to the theory, genes alone do not dictate phenotype. Rather, the potential for phenotypes that are typically male, female, or androgynous accompanies both XX and XY genotypes. Whether development proceeds along a classically male versus female continuum of traits depends upon relative concentrations of sex hormones during gestation and post-natal life.

Essentially, GTC theory argues that common traits, from sex-role attitudes to cognitive ability, are influenced by an individual's sex hormones. Both sexes produce a combination of male and female sex hormones (Nyborg, 1994). GTC theory considers the relative balance between the two types to be the key determinant of sexual differentiation. As used here, sexual differentiation can be thought of as the degree to which an individual manifests quintessentially masculine or feminine traits.

For females, GTC theory predicts that the relative level of estrogens or the estrogen-testosterone balance contributes to inter-individual variation in the following characteristics:

* height;
* ratio of fat to muscle;
* maturational tempo;
* sex-role identification;
* sociability;
* number of children;
* energy level;
* life expectancy; and
* performance and verbal IQ.

Specifically, GTC theory predicts that females having the highest levels of cognitive ability will be taller, leaner, more physically active, and less extroverted. Owing to lesser estrogen exposure and/or a higher ratio of testosterone to estrogens, these females are expected to mature later than their lower-ability peers; have less maternal orientation and fewer children; and be more career-oriented. The theory also anticipates that the highest-ability females will enjoy the longest life expectancies.

Table 1 summarizes the theory's predictions for females. The choice of five estrotypes--as opposed to some other number--is arbitrary (Nyborg, 1994). Under this typing system, females classified as E5 have the highest plasma levels of estradiol, a potent estrogen. As a consequence, the E5 female is the most sexually differentiated. E1 females, by contrast, have the lowest estradiol levels and the least feminized set of traits.

This dissertation took GTC theory a step further. It considered whether the life histories of mathematically and mechanically gifted females suggested them to be the least sexually differentiated of all. Put another way, it investigated whether females with atypical cognitive abilities differed from bright women with more typical talents. If so, a female with this atypical pattern might be better described not as an E1, but as an E0.

Prior to this study, neither GTC theory, nor the adapted model used here, had been tested in a female population. By contrast, the theory's predictions for males had been investigated in a sample of U.S. servicemen. The relative levels of androgens (and/or their metabolites) in these men were expected to influence an array of cognitive, somatic, and behavioral traits. When tested against life history data for more than 3,000 male veterans, the predictions were largely supported (Nyborg, 1994).

Research Questions

As noted, GTC theory predicts that highly intelligent females will be less sexually differentiated than females of lower ability. The adapted model anticipated that mathematically and mechanically gifted females would differ from other academically able women in a manner suggestive of an even lower degree of female sexual differentiation. The adapted model is summarized pictorially in Figure 1. Its validity was the key question in this investigation.

The adapted model made the following predictions derived directly from GTC theory.

1. *Height* will be greater in females with high levels of both mathematical and mechanical ability (HMHM females) than in controls.
2. *Body fat* will be lower in HMHM females than in controls.

Table 1
GTC Theory Predictions for Females

	Body Measures								Abilities		
Estrotype	Body Height	Fat/Muscle Ratio	Maturational Tempo	Sexual Identity	Sociability	No. of Children	Life Expectancy	Energy Level	Verbal	Performance	g
E5	--	++/--	++	++F	++	++	--	--	--	--	--
E4	-	+/-	+	+F	+	+	-	-	+	--	-
E3	Mean	Mean	Mean	F	Mean	Mean	Mean	Mean	Mean	Mean	Mean
E2	+	-/+	-	-F (M)	-	-	+	+	(-)	+	(+)
E1	++	--/++	--	--F (+M)	--	--	++	++	++	++	++

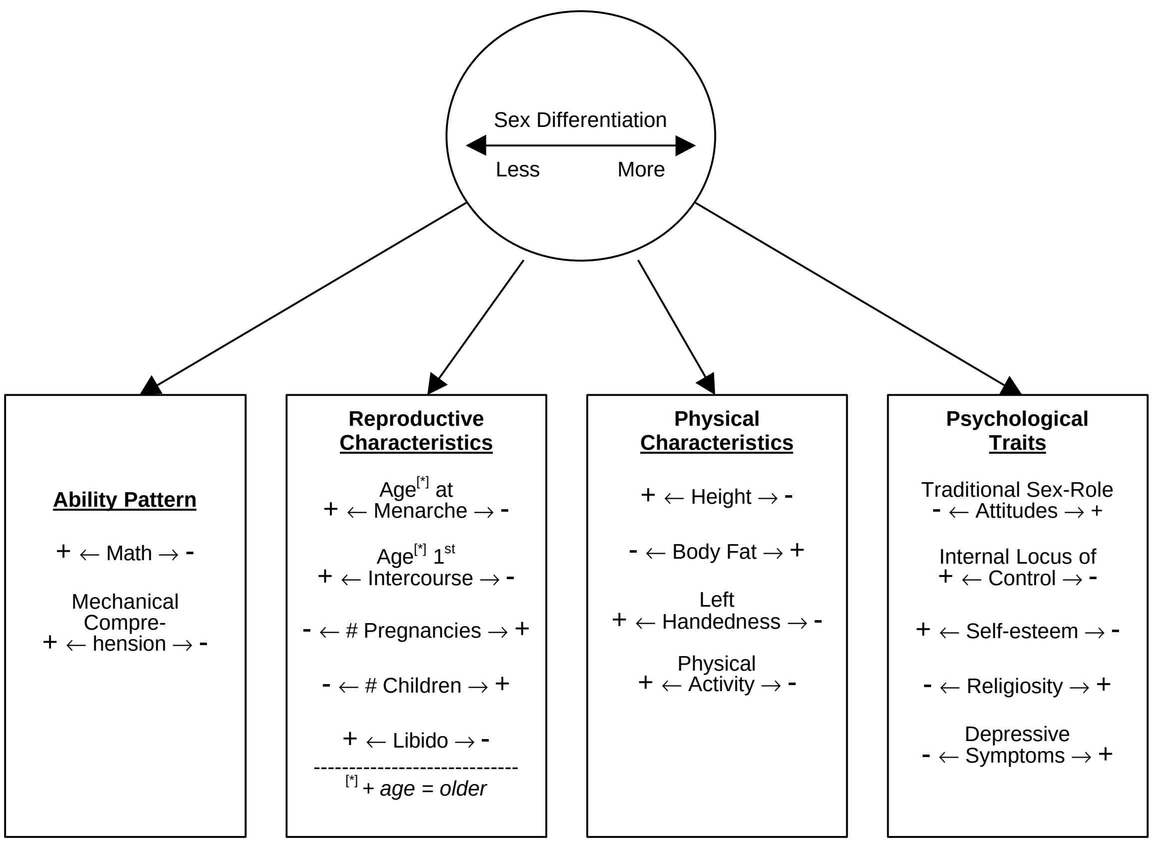


Figure 1. The dissertation model as adapted from GTC theory.

3. *Maturational tempo* will be slower in HMHM females than in controls.
4. *Physical activity* will be greater among HMHM females than among controls.
5. *Traditional sex-role attitudes* will be less pronounced among HMHM females than among controls.
6. *Parity* will be lower among HMHM females than among controls.

In addition, the adapted model included additional characteristics not explicitly cited by GTC theory. Two of these--libido and number of pregnancies--are logical extensions of Nyborg's premise that sexual differentiation influences reproductive history. For example, males generally have stronger sex drives than females (Jones & Barlow, 1990; Wilson & Lang, 1981). Accordingly, females very low in sexual differentiation would trend toward higher libido than more feminine women (see Figure 1). Chapter 2 reviews evidence that certain additional elements of the adapted model are potential indicators of sexual differentiation.

Expansion of the model led to adoption of seven additional predictions. These were as follows:

7. *Number of pregnancies* will be lower among HMHM females than among controls.
8. *Libido* in adulthood will be higher among HMHM females than among controls.
9. *Religiosity* will be lower among HMHM females than among controls.
10. *Locus of control* will be more internal among HMHM females than among controls.
11. *Left-handedness* will be more common among HMHM females than among controls.
12. *Depressive symptoms* will be less common among HMHM females than among controls.
13. *Self-esteem* will be higher among HMHM females than among controls.

Operational Definitions

The following definitions were used throughout this investigation.

* *Academically able* describes an individual whose score on the Armed Forces Qualification Test (AFQT) is at or above the 75^{th} percentile. The AFQT score is derived from the results of the Armed Services Vocational Aptitude Battery (ASVAB). Subjects sat for this examination in 1980.
* *Body fat* refers to a subject's body mass index (BMI). This measure is calculated by dividing weight in pounds by the square of height in inches, and multiplying the result by 703.1. The BMI is recognized as a far better indicator of adiposity than weight alone (Pi-Sunyer, 1988). Weight data are based on self-reports provided in 1981, 1986, 1990, and 1996.
* *Depressive symptoms* refer to scores on the Center for Epidemiologic Studies-Depression (CES-D) rating scale. The CES-D is a widely used self-report instrument designed to measure depressive symptoms in the general population (Radloff, 1977). Subjects completed the scale in 1992.
* *Handedness* refers to responses to the question, were you born naturally right-handed or left-handed? It was posed in 1992.
* *Height* refers to self-described adult height in inches, as reported in 1985.
* *High-mathematical, high-mechanical (HMHM) ability pattern* describes an outcome on the Armed Services Vocational Aptitude Battery that places a subject among the top decile of all test-takers (male and female) on subtests of mathematics knowledge, arithmetic reasoning, and mechanical comprehension.
* *Libido* is represented by the response given when asked, how many times did you have sexual intercourse during the past month? While less sophisticated than a standardized instrument, this approach has a history of use in assessing sexual frequency (Seidman & Rieder, 1994). The question was posed in 1984 and 1985.
* *Locus of control* is a participant's score on the Pearlin Mastery scale. This standardized instrument was administered in 1992. It measures an individual's sense of control over life events (Pearlin, Lieberman, Menaghan, & Mullen, 1981).
* *Maturational tempo* is represented by two variables. The first is the subject's recalled age at menarche, a commonly used marker for measuring age at sexual maturity (Hassler, 1991). The second is recalled age at first intercourse. This variable shows a relationship to age at menarche (Andersson-Ellstrom et al., 1996). Accordingly, it offers a second measure of sexual maturation.
* *Number of pregnancies* refers to the total number of pregnancies reported by subjects in various survey years. This variable includes all pregnancies, including those terminated by spontaneous or elective abortion.
* *Parity* is the cumulative number of live births reported during various survey years.
* *Physical activity* refers to responses regarding participation in high school athletics or cheerleading. Although arguably an imprecise measure, medical studies frequently rely on a simple inquiry of this

type to assess degree of physical energy expenditure (Ainsworth, Sternfeld, Slattery, Daguise, & Zahm, 1998; Mittendorf et al., 1995).

* *Religiosity* refers to a participant's self-reported frequency of attending services at a church or other house of worship. This is a commonly used indicator for assessing this trait (Balakrishnan & Chen, 1990; Miller & Hoffman, 1995).

* *Self-esteem* is the score on the Rosenberg Self-Esteem Scale. This a standard instrument for measurement of self-concept (Rosenberg, 1965).

* *Traditional sex-role attitude* is represented by a variety of expressed values and behaviors. These include support for a lifestyle characterized by male participation in the workforce and female commitment to nurturant, home-centered activity. Such attitudes are assessed by scores on the Scale of Traditional Attitudes, an inventory that measures attitudes toward female participation in the workforce (CHRR, 1997). Desire to marry, expected age at marriage, number of children wanted, and occupational choice are also used to assess this trait. In the interest of brevity, it is hereinafter termed *traditionality.*

Chapter 3 includes detailed discussion of the Armed Services Vocational Aptitude Battery and the psychological rating scales used in this research.

Limitations

In the broadest sense, this study asked whether hormonal factors influence human development to a degree heretofore unappreciated by many scholars in the field. Such an inquiry--grounded in GTC theory--was ripe for investigation.

Nonetheless, the data imposed certain limitations. The most obvious is that sexual differentiation was assessed via surrogate markers, rather than actual hormone levels. The latter were not available for the subjects in this study, nor is any existing data set known to contain such information (Nyborg, 1994).

In addition, several factors compromised the statistical power of the study. Subjects were selected from the National Longitudinal Survey of Youth (NLSY). Due to oversampling of disadvantaged populations, persons in the upper quartile of ability were statistically underrepresented in the NLSY. As a consequence, sample sizes for both focal and control groups were smaller than would be expected from a pool of 6,000 females.

Budget constraints forced downsizing of the NLSY sample in 1985 and again in 1994. Accordingly, for variables measured after these events, the number of subjects in each ability group was reduced further. A liberalized alpha was adopted in an effort to compensate for these reductions in statistical power. In this context, the overall pattern of results--rather than the results relative to any one variable--provides the most meaningful basis for analysis.

Of course, this investigation did not (and could not) examine *all* potential associations or influences related to the HMHM ability pattern. Accordingly, the relationships discussed here do not represent its only correlates. Genetic, social, and environmental factors almost certainly influence the study variables. The lack of attention to such factors in this investigation was not intended to suggest their irrelevance.

Chapter 2: Review of the Literature

Research into the many facets of cognitive development has created a vast literature on the distribution and correlates of general intelligence. Comprehensive reviews and analyses of these findings are readily available (Gardner et al., 1996; Jensen, 1998; Storfer, 1990).

Investigations that attempt to explain variation in the performance of females on standardized measures of mathematics-related competencies comprise but a minute fraction of the literature in cognitive psychology. However, literature on between-sex differences also provides some basis for evaluating likely sources of within-sex variation.

This concept can be appreciated by considering a trait such as running speed. Clearly, height plays a role in between-sex differences (Wilson, 1992). Its applicability to understanding within-sex differences is equally obvious. Accordingly, a portion of this review will consider the literature on between-sex differences in mathematical performance.

Noticeably absent from this review will be findings regarding sex differences in mechanical reasoning ability. The literature concerning this aptitude is extremely small. Primarily, it addresses the measurement of ability and teaching considerations. Efforts to explain the large sex difference in this aptitude or determine its correlates in female populations are surprisingly absent.

An exception is a recent report confirming the large sex difference in spatial-mechanical reasoning ability (Hamilton, Nussbaum, Kupermintz, & Kerkhoven, 1995). This large-scale analysis found a correlation between this ability and visiting science museums, but no association with instructional variables. Because mechanical and spatial abilities are considered to be closely related (Vernon, 1950), notable findings about sex differences in the latter will be included in this review.

The relevant literature generally examines variation in performance as a function of social, psychological, or biological characteristics. These classifications can overlap; ethnicity, for instance, has both socio-cultural and biological dimensions. In the interest of simplicity, however, each factor will be classified under a single category.

Social Models of Between-Sex Differences

Among many social scientists and educators, social factors are considered the sine qua non for explaining virtually all sex differences (Nyborg, 1994; Tooby & Cosmides, 1989). Obvious exceptions to this general rule are pregnancy, lactation, menstruation, and ejaculation. These four functions are the only characteristics subject to absolute limitation based on sex (Nyborg, 1994). All other human faculties can be found in both sexes, although sex-typical patterns have long been recognized.

This theoretical outlook sees human beings having no particular nature or proclivity to develop certain temperaments or abilities at birth. Essentially, it holds that the human mind is a blank slate or tabula rasa, waiting to be molded by its environment. So pervasive is this view that it has been called the Standard Social Science Model, or SSSM (Tooby & Cosmides, 1989). Essentially, the SSSM embraces the Durkheimian mandate to explain social facts with other social facts. Accordingly, differences in temperament, values, and cognitive outcomes are attributed to environmental influences (Nyborg, 1994).

Social facts often cited to explain lower female achievement in mathematical and spatial reasoning include societal and parental stereotyping of sex roles (Jeffe, 1995; Meece & Jones, 1996; Parsons, Adler, & Kaczala, 1982; Tocci & Englehard, 1991; Wilder, 1996). Unequal attention from teachers is another. Girls allegedly receive less feedback from instructors than boys, which, in turn, is presumed to erode their self-confidence, interest, and abilities in male domains such as math (AAUW, 1989; AAUW, 1991; Becker, 1981; Jones, 1989; Midgley, Feldlaufer, & Eccles, 1989; Sadker & Sadker, 1986; Wilson, 1980)

Similarly, socialization theorists fault teaching materials. They argue that female performance suffers from sexist curricula such as textbooks that contain more pictures of boys than girls or that portray the girls in typically female endeavors (CRW, 1996). Teaching methods are also criticized--particularly so-called competitive learning, which is alleged to better serve the learning needs of males than females (CRW, 1996).

The ability of these arguments to explain between-sex variation in female mathematical achievement is debatable. One concern relates to the very quality of the facts central to these arguments. Questions have been raised about the validity of claims that teachers preferentially attend to boys. Much of the data have yet to be published in peer-reviewed literature (Sommers, 1994).

An experiment designed to test this allegation was recently published in a scholarly journal (Altermatt, Jovanovic, & Perry, 1998). Researchers studied teacher responses to students in six science classrooms. In three, teachers did call on boys more frequently. However, analysis showed that the differential pattern resulted from a tendency of the boys to volunteer answers more frequently than girls did.

In addition, those who charge that teachers preferentially attend to boys often classify contrasting types of feedback (both compliments and reprimands, for instance) under the single category of teacher attention. There is no consensus that such qualitatively different responses produce similar effects.

The argument that teacher behavior contributes to sex differences in mathematical achievement is further weakened by findings comparing test scores between single-sex and coeducational institutions. Studies to date have yet to find differences in female mathematics performance attributable to school type (AAUWEF, 1998; Bell, 1989; Finn, 1980; Steedman, 1985). Obviously, if preferential treatment of boys by teachers contributed to differences in achievement, girls in single-sex schools would outperform those attending coeducational institutions.

The charge that parents socialize girls to see themselves as less competent in mathematical domains also falters under careful analysis. A comprehensive review of almost 200 studies of sex-related child-rearing practices included numerous investigations related to cognitive development (Lytton & Romney, 1991). Among North American parents, the effect size for encouragement of achievement was a negligible .02 (favoring males). A larger effect favoring males (d = 0.29) was found for studies from other Western countries.

Effect sizes for three studies regarding parental encouragement of math achievement were .08, .00, and -.12. These outcomes are inconsistent with claims of differential treatment by parents. Similarly, analysis of parental involvement in the education of mathematically and verbally precocious adolescents found no support for the notion that parents encourage girls in verbal areas and boys in mathematics-related domains (Raymond & Benbow, 1986). Many girls do take a considerable amount of mathematics coursework in high school (Benbow, 1988). This is difficult to reconcile with charges that educators and parents teach girls not to attempt achievement in math.

Perhaps the most serious shortcoming of socialization theories is their failure to account for the temporal nature of sex differences in performance. As discussed earlier, sex differences in mathematics performance favoring males either do not emerge until puberty or become more pronounced around that time.

By admission of those who support the arguments cited above, these alleged influences--sexist textbooks, unequal treatment by teachers, and social stereotypes--are as present in elementary school as during the high school years (CRW, 1996; Sadker & Sadker, 1986). Yet, throughout most of the elementary grades, no correlation with female performance can be shown. As noted earlier, most studies find that girls perform as well or better than boys on standardized tests of mathematics prior to puberty or its onset (Willingham & Cole, 1997).

Accordingly, it appears that despite a tendency to be more compliant with the wishes of adults during the first decade of life, girls routinely defy allegedly sexist role models presented to them during the primary grades. No one has explained how these influences could simultaneously have no effect or positive effects on female mathematical performance prior to adolescence, only to cause a rather sudden reversal of fortune around the time of puberty.

Arguably, an explanation that indirectly addresses this set of facts is the gender intensification theory (Hill & Lynch, 1983). The concept holds that pressures for sex role conformity increase at adolescence. Halpern explained: (1992, p. 200)

> Adolescence is a transitional period of life during which children undergo a metamorphosis from which they emerge as adults. Peer interactions are especially important, and there are strict sanctions against sex role inappropriate behavior. The need to be just like everyone else is strong. Strict conformance to sex role stereotypes would also require boys and girls to conform to sex-typed cognitive activities (e.g. avoiding mathematics and science coursework for girls and avoiding poetry and literature for boys).

Certain facts, however, render this theory unsuitable as an explanation for the shift in female performance that occurs around the time of puberty. If females feel pressure to underperform boys in mathematics, one would expect them to take schoolwork in this subject less seriously and receive lower grades. Yet, despite poorer performance on standardized tests, girls achieve higher grades than boys in math throughout their high school years (CRW, 1996; Willingham & Cole, 1997). Grades are far more visible to peers than test scores. This fact alone seriously undermines the theory.

This notion fails to account for other important aspects of cognitive sex differences. One underappreciated, but relevant fact is that male superiority does not hold for all aspects of mathematics. Rather, it is found in the domains of mathematical reasoning and geometry (Willingham & Cole, 1997). Throughout life, females demonstrate better computational skills than males. Composite scores on tests given in grades 4, 8, and 12--as well as superior female performance on the numerical operations subtest of the Armed Services Vocational Aptitude Battery--provide ample documentation of this phenomenon (Bock & Moore, 1986; Willingham & Cole, 1997).

Another investigation shows a similar pattern of female superiority on certain types of questions on the ACT--a college admissions test (Doolittle & Cleary, 1987). This analysis credits females with better skill at answering questions requiring algorithmic or computational ability. These findings are difficult to reconcile with suggestions that socialization pressures cause females to underachieve in math from adolescence onward.

One must also ask whether the magnitude of female decline in mathematical reasoning ability is consistent with the notion of gender intensification. The term implies an effect of considerable size--and may even prompt expectation of a bimodal ability distribution in sex-typed domains. The reader must decide whether the actual change in female performance--from a standard mean difference of 0.03 in their favor at Grade 4 to one favoring boys by about .07 in Grade 12 (Willingham & Cole, 1997)--can reasonably be termed intensification toward a sharply delineated sex role.

Finally, socialization models must be weighed in the context of the secular trend in scores on the Scholastic Aptitude Test (recently renamed the Scholastic Assessment Test). Figure 2 is adapted from data provided by the test developer (ETS, 1998). As it illustrates, the ratio of female-to-male scores on the mathematics section has remained almost constant, at about 92%, for about three decades. If social factors explain this phenomenon, their influence is exerted in a manner that allows females to calibrate the magnitude of their lesser performance with extraordinary precision--and hold it constant through decades of major changes in sex roles. The implausibility of such a scenario seems obvious.

Social Models of Within-Sex Difference

Focusing these same social models on within-sex differences results in a somewhat different picture. Regardless of one's view of social models of between-sex differences, few of these alleged influences can explain within-sex differences. The reason is obvious. High-performing girls generally experience the same societal messages as poor performers.

For example, sexist textbooks have long been suspected of discouraging female math achievement (CRW, 1996; Hoffman, 1982). Yet, girls who do well on standardized math tests often read the same textbooks as those who fare poorly. Both high and low scorers experience similar societal messages regarding sex roles. Some may have classes where teachers give preferential treatment to males. Yet, a considerable number of girls outscore many of the boys in their classes or choose to take higher-level mathematics courses in high school (Benbow, 1988; Willingham & Cole, 1997). That females as a group have not achieved parity with males in mathematics-related domains has not deterred some girls from pursuing advanced coursework--nor from reaching substantial levels of achievement.

While all females experience certain similar influences, some social factors obviously vary among individuals. Family background is an obvious example. Some parents undoubtedly encourage educational achievement more than do others.

Researchers have published numerous studies regarding the family backgrounds of academically gifted children.

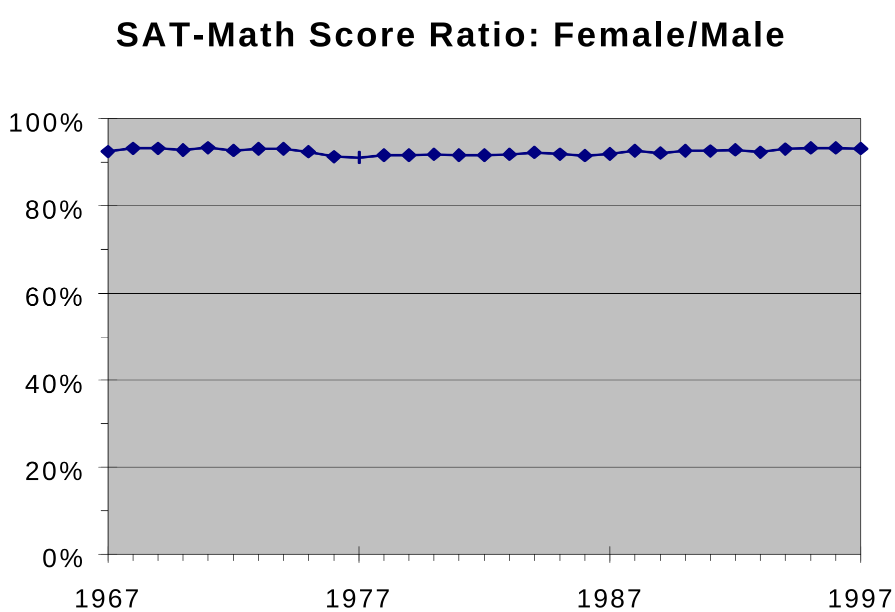


Figure 2. Sex differences on the math SAT.

Mean female score on the mathematics section of the SAT as a percentage of the male score during the past three decades.

Information specific to the home lives of females who excel in mathematics and related domains is quite limited. Nonetheless, several factors deserve mention.

Ethnicity. Within the United States, a disproportionate number of mathematically gifted females are of Asian ethnicity (Moore & Stanley, 1987). Asian ancestry correlates with performance not only at the high end of the ability curve, but also at the mean. In 1997, the average SAT-math score of Asian-American females considerably exceeded that of White females and trailed that of White males by a mere two points (ETS, 1998). That these females perform as well as White males contradicts assertions that sex-role stereotypes and educational practices in the U.S. impede the learning potential of girls.

Parental encouragement, tutoring, and expectations are among factors cited to explain the superior performance of Asian students both in the U.S. and abroad (Moore & Stanley, 1987; Rowe, 1994; Wong, 1992). A contributing role for any (or all) of these factors is certainly plausible. The more vexing question is whether these influences fully account for the superior performance of Asians over other ethnic groups.

Two factors argue against such a purely environmental explanation. The first is the tendency of Asian females to outperform their peers of other ethnic backgrounds on measures of spatial ability (Storfer, 1990). Spatial skills are rarely part of any academic curriculum (Bock & Moore, 1986). As a consequence, parental encouragement and help with schoolwork make a weak explanation for the high ability of Asian females in this arena.

This raises the possibility that differences in cognitive structures and processes may underlie disparities in spatial ability. In light of the longstanding view that spatial skills influence some aspects of mathematical performance (Bock & Moore, 1986; Stanley, 1993; Storfer, 1990), one cannot rule out the possibility that biological differences similarly contribute to the higher mathematical performance of Asian females. If this is so, the significance of ethnicity may derive more from biological influences than from child-rearing practices.

An equally important consideration is outcome in verbal domains. The familial and personal traits credited with giving Asian students a strong edge in mathematics convey no similar advantage in the verbal arena. To the contrary, females of Asian descent consistently underperform their White counterparts on tests of verbal ability (Storfer, 1990). For example, White females outscored Asian females by 30 points on the 1997 SAT-verbal (ETS, 1998). This represents an order of magnitude almost identical to the advantage shown by Asian females on the SAT-math. This dichotomy also challenges the notion that the family environment alone explains the relationship between Asian ethnicity and mathematical achievement.

In addition to Asian females, those of Jewish descent are also statistically overrepresented among the most mathematically talented (Storfer, 1990). However, high spatial ability tends to be absent in these females--as well as in their male counterparts (Storfer, 1990). Explaining why Jews and Asians nonetheless perform similarly in the realm of general mathematical ability has been a challenge for developmental scientists.

Studies of Jewish populations have revealed an unusually high level of verbal reasoning ability (MacDonald, 1994). This marked verbal aptitude appears to best explain the overrepresentation of Jews among persons of high mathematical ability (Storfer, 1990). Strategies based in verbal reasoning appear useful for solving many mathematical problems. Those based in spatial reasoning, however, may be more efficient (Stanley, 1993).

Parental characteristics. The correlation between parental factors and female mathematical achievement is not limited to Asian families. Presence of both parents in the household or frequent contact with the father also correlate with mathematical achievement (Belz & Geary, 1984; Campbell, 1996; Sandqvist, 1995; Storfer, 1990).

Both parental intelligence and father's occupation have been associated with performance on standardized tests (Belz & Geary, 1984). A study that examined the family backgrounds of precocious adolescents found both parents tended to be unusually intelligent (Benbow, Zonderman, & Stanley, 1983). This is consistent with the principle of assortative mating-- the tendency for individuals to marry persons of similar traits. Among the parents of these unusually gifted children, the magnitude of assortative mating for intelligence was significantly higher than among the general population.

Socioeconomic status. Within the U.S. population, economic disadvantage clearly correlates with poorer cognitive performance--both generally and in mathematical domains (Bock & Moore, 1986; Entwisle & Alexander, 1996; Jensen, 1998). In psychosocial terms, the factor that best accounts for this finding is the quality of the residential environment.

Children from lower socioeconomic classes often live in homes providing few sources of cognitive stimulation, such as books, to engage their latent abilities (Storfer, 1990). A significant correlation between quality of the home environment and intellectual performance in preschool and elementary grades has been documented (Baker, Keck, Mott, & Quinlan, 1993; Crane, 1996).

School quality, on the other hand, explains the correlation less well. In a comprehensive review, Ceci concluded that virtually all schools in the U.S. provide sufficient stimulation to allow individuals to develop their cognitive abilities. Accordingly, the reported relationship between years of education and intelligence scores may be explained by continued presence in *any* school, rather than by attendance in those noted for the achievement of their students (Ceci, 1991). Although arguably a minor point, some of history's greatest mathematical minds--Isaac Newton and Michael Faraday among them--did not come from privileged homes nor enjoy access to good schools in childhood. Faraday, in fact, had almost no schooling at all (Wilson, 1992).

However strong this correlation, its causal significance remains unclear. Contrary evidence comes from several venues, including cross-cultural comparisons. On tests of mathematical ability, the Chinese outperform Caucasians living in far more affluent circumstances (Geary, Bow-Thomas, Liu, & Siegler, 1996; Lynn, 1987; Stevenson et al., 1990). Similarly, socioeconomic disadvantage is by no means a bar to mathematical achievement in the U.S. In a group of more than one thousand mathematically gifted high school students, about one in six came from homes of below-average socioeconomic status (Lubinski & Humphreys, 1990).

Studies of adopted children also challenge the notion that economic advantage is causally related to ability. When infants born to low-income parents are adopted into middle-class homes, they do not develop the degree of intellectual advantage that would be expected were economic status a major influence on cognitive performance (Locurto, 1990; Rowe, 1994).

In the elementary grades, such children have scored better on intelligence tests than would be predicted based on the background of their biological parents. Nonetheless, their cognitive performance has failed to approach that shown by the natural children of the adoptive parents (Weinberg, Scarr, & Waldman, 1992). Most disappointing of all, higher-than-expected test scores seen in elementary grades have not been sustained--and have all but disappeared by age 17. The results are unpleasantly reminiscent of findings showing the intellectual benefits of compensatory education to be transient in many youth (Currie & Thomas, 1995; Holden, 1990).

These data appear destined to fuel continued debate as to whether socioeconomic status is a causative factor in ability--or simply, as some argue, a marker for its underlying factors (Jensen, 1998). To the extent that socioeconomic status is a marker for the genetic endowment that children receive from their parents, the difficulty of separating correlation from causation is evident.

Brodnick and Ree have made a first attempt to make this distinction. They modeled academic achievement in a sample of 340 college students. After controlling for general ability, socioeconomic status did not account for any of the remaining variance in outcome (Brodnick & Ree, 1995).

To summarize, evidence that social factors account for between-sex or within-sex variation in mathematical ability is minimal. Family background, however, is an exception. While correlations are compelling, their causality is difficult to determine. Putative influences such as the quality of the home environment or parental involvement in a child's education may themselves be dependent variables--a consequence of parental temperament, education, and general intelligence. On the other hand, these correlates may represent factors that are primarily markers for independent variables yet have some causal significance of their own.

Nothing herein is intended to deny the intuitive appeal of explanations that attribute cognitive outcomes to socioeconomic status and/or the social environment. Widespread allegiance to these concepts understandably derives from a sense that these factors are readily amenable to intervention. Scientists, however, are obliged to recognize that good intentions are insufficient to establish what is true--and that the appeal of a concept does not covary with its sustainability in the scholarly arena.

Psychological Profiles

A number of investigators have examined the personality traits of females who excel in mathematics. Noteworthy findings include:

* a cognitive style favoring thinking over feeling (Mills, 1998);
* a negative correlation with social or religious interests (Stanley, 1993);
* an internal locus of control (Brody & Benbow, 1986; Starr, 1979);
* masculine interests or self-concept (Horn & Turner, 1974; Lippa, 1998a; Selkow, 1987; Signorella & Jamison, 1986)
* high achievement motivation relative to other females (Mills, 1998);
* lower interest in interpersonal matters as compared to other females (Mills, 1998);

Results regarding self-esteem have been mixed. Some investigators report higher self-esteem among mathematically gifted females (Starr, 1979). Others have been unable to document differences between mathematically talented females and controls (Brody & Benbow, 1986).

At first glance, the above findings may seem largely unrelated. To the contrary, there is a pattern here--a rather striking one at that. In essence, it is a profile more consistent with the psychological characteristics of males than females. For example:

* The cognitive style of an estimated two-thirds of U.S. males favors thinking over feeling (Hammer &

Mitchell, 1996). By contrast, about 60% of women prefer feeling over thinking.

* Males show less religiosity than females (Levin, Taylor, & Chatters, 1994; Miller & Hoffman, 1995).
* Males typically have a more internal locus of control than females (deBrabander & Boone, 1990; Kunhikrishnan & Manikandan, 1992; Lopez & Staszkiewicz, 1985; Nunn, 1994).
* Males report fewer career interests classified as social than do females (Holland, 1972; Kaufman & McLean, 1998).
* Some divergence in self-esteem favoring males emerges during the college years (Maccoby & Jacklin, 1974), and recent data suggest that males may, in fact, have somewhat higher self-esteem than females (Dukes & Martinez, 1994; Richman, Clark, & Brown, 1985).

For behavioral scientists, the challenge is to understand why some females develop traits atypical of their sex. As already discussed, general societal messages cannot be responsible.

Parental socialization, by contrast, will differ from one home to the next. Accordingly, some parents may behave in ways that encourage their daughters to develop a more masculine personality. However, the notion that such parental practices occur with much frequency seems questionable. More importantly, evidence to support such a concept is lacking.

In their comprehensive analysis of parental practices in North America, Lytton and Romney (1991) found only one area in which both mothers and fathers differentially socialized their offspring depending on the child's sex. That area was the encouragement of sex-typed activities (e.g., doll-play for girls). This suggests that atypical interests may develop despite parental efforts to encourage more sex-typed traits.

In short, the psychological data indicate that females who excel in mathematics are less stereotypically feminine than their peers. Societal influences cannot logically account for this phenomenon. Data that demonstrate an influence of differential socialization by parents are similarly lacking. If only by default, these findings lend credence to the possibility that biological influences are at work.

The Biological Context

Knowledge about the biological correlates of mathematical ability is limited. By contrast, considerable literature exists regarding the biological aspects of general intelligence. Several books offer comprehensive reviews (Jensen, 1998; Vernon, 1993).

A few investigators have studied mathematical ability apart from general intelligence. Particularly intriguing is a report that physical and medical well-being correlate more highly with mathematical giftedness than with economic privilege (Lubinski & Humphreys, 1992). In this comparison, adolescents who scored in the highest 1% of mathematical ability enjoyed significantly better health than peers ranked in the top 1% on an index of socioeconomic status. The measurement of health status relied on 43 indicators of well-being, such as frequency of doctor visits or illnesses requiring bed rest.

Nonetheless, certain health problems are especially common among the mathematically gifted. While the association between myopia and general ability is well-established (Benbow, 1986; Cohn, Cohn, & Jensen, 1988; Miller, 1992), one report suggests that the relationship may be even stronger for mathematical giftedness. Karlsson studied visual acuity of individuals who had achieved highest honors at an Icelandic university. Among females, myopia was almost twice as common among those who graduated in mathematics than in the humanities (Karlsson, 1976). In addition to their tendency toward myopia, both mathematically and verbally precocious youth are unusually prone to allergies (Benbow, 1986; Benbow, 1988).

While rarely considered a health problem, left-handedness also correlates with mathematical ability. Among young adolescents with exceptional mathematical or verbal reasoning ability, about 15% were left-handed--a rate about 50% higher than that of controls (Benbow, 1986). Some have confirmed a finding between left-handedness and mathematical aptitude or involvement in certain math-intensive disciplines (Smith, Meyers, & Kline, 1989; Temple, 1990). Other investigations have not (Peters, 1991).

A relationship between spatial ability and various degrees of left-hand preference has also been reported (Annett, 1992; Burnett, Lane, & Dratt, 1982; Natsopoulos, Kiosseoglou, & Xeromeritou, 1992). Not all investigations have found such a relationship, however (Martino & Winner, 1995).

Work by Casey and colleagues may help explain the negative findings. These investigators have examined the interaction between handedness, family history, and mental rotation ability. Among female college students majoring in math and science, those most likely to be competitive with males on spatial tasks were right-handers with non-righthanded relatives (Casey & Brabeck, 1989). A follow-up investigation among minority students enrolled in a high school science and technology program yielded similar results. Right-handed girls with non-righthanded relatives outperformed left-handed girls and right-handed girls with similarly handed relatives (Casey, Colon, & Goris, 1992a).

Obviously, further study will be needed to sort out these conflicting data. That left-handedness is more common among males than females is nonetheless clear (Kurian & Santhakumari, 1994; le Roux, 1979). Among adolescents with unusually strong mathematical or verbal reasoning ability, significantly more males than females report left-hand preference as well (Benbow, 1988).

Research has found no difference in general intelligence between left-handers and right-handers (Hardyck, Petrinovich, & Goldman, 1976). Thus, these findings suggest that left-handedness, or a family history of it, may be a marker for particular types of cognitive traits rather than of general ability per se (Benbow, 1988).

Genetic factors influence both handedness and myopia (Annett, 1979; Boklage, 1981; Cohn et al., 1988; Lin & Chen, 1987; Teikari, Kaprio, Koskenvuo, & Vannas, 1988). The heritability of mathematics-related competencies similarly appears to be considerable. Mathematical aptitude is more similar among identical twins than fraternal twins (Lange & Fischbein, 1990). The greater genetic similarity of identical twins is the most obvious explanation for this finding.

While few have estimated the heritability of mathematical ability as a characteristic distinct from general intelligence, one investigator reports that heritable factors account for about 60% of the variation in mathematical performance (Gillis, DeFries, & Fulker, 1992). Estimates of the heritability of spatial competence range from 32 to 63% (Bratko, 1996; McClearn, Johansson, Berg, Pedersen, & et al., 1997).

The specific mechanisms of heritability remain to be elucidated. Work continues on efforts to identify specific genes that may contribute to high levels of various abilities (Plomin & Thompson, 1993). While these investigations may eventually untangle the mystery of cognitive differences, the most compelling data to date come not from the genetics laboratory, but from research regarding the effects of hormones on cognition.

This area of inquiry takes its cue from the puberty-related onset of sex differences. It considers the possibility that puberty brings physiological changes that affect not only characteristics such as voice and appearance, but also the structure and function of the brain. The implication is that puberty is accompanied by changes in human biology beyond that which is readily seen and heard.

The empirical data that support this concept are considerable. Advances in the neurosciences make clear that some of the most profound differences between males and females lie in the brain. This 3-pound organ controls cognitive processing, and, ultimately, all behavior (Andreason, Zametkin, Guo, Baldwin, & Cohen, 1994; Mead & Hampson, 1996; Nyborg, 1994; Wilson, 1992).

The relevant differences are not necessarily present at birth. In both man and animals, certain sex-linked traits develop over the course of time. Contributing to their emergence are hormones--specifically those related to sexual differentiation (Beatty, 1984; Hampson, 1995; Moffat & Hampson, 1996; Nyborg, 1983; VanGoozen, 1994; VanGoozen, Cohen-Kettenis, Gooren, Frijda, & VanDePol, 1995).

Evidence indicates that hormones influence not only sex-typical behaviors such as grooming and aggression, but also navigational skills (Barr, Gibbons, & Moyer, 1976; Bronstein, Wolkoff, & Levine, 1975; Butovskaya & Kozintsev, 1996; Moore, 1986; Thor, Harrison, Schneider, & Carr, 1988). In both humans and many animal species, navigational skills are superior in males (Geary, 1995). Conversely, where species-specific conditions impose more spatial demands on females than males, the former show better navigational skills (Geary, 1998a). Such findings lend support to arguments that other cognitive sex differences also have biological underpinnings (Geary, 1998a; VanGoozen, 1994).

Hormones affect the organization of the brain as well as activation of neural processes (Nyborg, 1983; Williams & Meck, 1991). Organizational effects have lifelong implications, while activational effects do not. Some evidence suggests that prenatal exposure to testosterone affects brain organization in a manner that increases the likelihood of left-handedness and immune disorders such as allergies (Benbow, 1988; Geschwind & Behan, 1982). As noted earlier, these are correlates of mathematics-related abilities.

The notion that hormonal mechanisms mediate some of the sex difference in mathematical ability is also consistent with the puberty-related decline in standardized test scores among females. The increase in the magnitude of sex differences in spatial abilities at puberty similarly conforms to such a theoretical model. Equally important to this empirical view is a sizable literature derived from the description and treatment of individuals whose endocrine status has been altered by genetic, developmental, or exogenous factors.

The list of conditions causing atypical hormone profiles is considerable. Some, such as Turner's syndrome in females, result from chromosomal abnormalities. As a consequence of having only one X-chromosome, females with this condition have a profound deficiency of estrogens. If left untreated, these individuals have extremely low spatial ability--along with a host of other traits related to body size, personality, and interests (Nyborg, 1984). Afflicted females also have difficulties with basic arithmetic. These may reflect deficits in retrieving facts for executing numerical operations (Rovet, Szekely, & Hockenberry, 1994).

By contrast, females afflicted with congenital adrenal hyperplasia (CAH) show a markedly different life history. Due to an inborn metabolic error, these individuals secrete unusually high levels of testosterone during gestation. Diagnosis can generally be made at birth, and abnormal hormone secretion brought into the normal range with pharmacotherapy.

Affected girls nonetheless retain distinct characteristics. They have high spatial ability for their sex (Resnick, Berenbaum, Gottesman, & Bouchard, 1986); express little interest in motherhood; and have an atypical interest in objects as opposed to people (Nyborg, 1984). As children, CAH females are often unusually active relative to unaffected girls. They are also more likely to be tomboys and to engage in rough-and-tumble play (Ehrhardt, Epstein, & Money, 1968).

The high levels of testosterone to which CAH girls are exposed in utero appear to explain these characteristics (Hoyenga, 1993). A relationship between testosterone levels in the amniotic fluid and spatial ability has been reported in a group of normal girls as well. At the age of 7, mental rotation ability in these girls correlated positively with their testosterone exposure at 16 weeks in utero (Grimshaw, Sitarenios, & Finegan, 1995).

Clinical research further supports an effect of hormones on personality and spatial ability. Estrogens have been shown to promote greater sociability and extroversion (Arushanyan & Borovkova, 1989; Halbreich, Lemus, Lieberman, Parry, & Schiavi, 1990; Herrmann & Beach, 1978; McQueen, Wilson, & Fink, 1997; Rodriguez & Grossberg, 1998). In addition, spatial ability fluctuates with the menstrual cycle. Ability is highest when levels of estrogens are low (Hampson, 1990a; Hampson, 1990b; Phillips & Silverman, 1997; Silverman & Phillips, 1993).

Other investigations have found that females perform better on tests of deductive reasoning during menstruation than while in the mid-luteal phase of their monthly cycles (Kimura & Hampson, 1993). As levels of both estrogens and progesterone are low during menses, this finding is consistent with an adverse effect of female hormones on this type of task. Reports that link later sexual maturation among females with better spatial ability also suggest a negative influence of estrogens on certain cognitive outcomes (Diamond, Carey, & Back, 1983; Ray, Newcombe, Semon, & Cole, 1981; Sanders & Soares, 1986; Waber, 1977).

Evidence for this phenomenon remains somewhat inconsistent (Bryden & Vrbamcoc, 1988; Geary & Gilger, 1989; Rierdan & Koff, 1984). Waber (1977) observed better spatial skill among slower maturing adolescents in an initial investigation, but found somewhat contradictory results in a follow-up study. Early and late maturing subjects did not differ when tested with a battery of common spatial tasks (Waber, Mann, Merola, & Moylan, 1985). However, late maturers did perform significantly better on the Rey-Osterrieth Complex Figure Copying test. This visuo-constructive task requires subjects to copy (freehand) a highly complex abstract design. Better performance indicates greater facility for mental visualization of complex objects and may be relevant to mechanical aptitude.

Findings from studies of male-to-female transsexuals also support a relationship between sex hormones and spatial ability. Ability declines in these individuals, who receive feminizing hormones as an adjunct to surgery (Hampson, 1990a; Hampson, 1990b; Phillips & Silverman, 1997; VanGoozen, 1994). Nonetheless, the relationship between estrogens and spatial ability does not appear to be linear, but U-shaped (Nyborg, 1994). Both very high and very low levels of estrogens appear to inhibit competence with spatial tasks.

In addition to estrogens, testosterone also appears to influence specific cognitive abilities, personality factors, and physical characteristics. Several investigators report a positive correlation between salivary testosterone and spatial ability in adult females (Gouchie & Kimura, 1991; Hassler & Nieschlag, 1989). These findings are relevant to the matter at hand because of the presumed relationship between spatial ability and mathematical performance as well as the overlap between spatial and mechanical aptitudes (Bock & Moore, 1986; Vernon, 1950).

Testosterone levels in females have also been correlated with a more masculine or androgynous personality (Baucom, Besch, & Callahan, 1985). As already discussed, mathematically talented females tend to have such a profile. Coupled with the findings reviewed thus far, these results constitute compelling evidence that hormones influence numerous aspects of personality, behavior, and cognition.

Chapter 3: Methods

Research Design

This study compared the life histories of two groups of academically able females. The focal group was comprised of females with high mathematical and mechanical (HMHM) reasoning ability. An approximately equal number of females were selected as controls.

Subjects in the control group were academically able, but lacked the particular strengths in mathematics and mechanical reasoning demonstrated by those in the HMHM group. The two groups were compared on somatic, reproductive, and personality variables representing the conceptual model shown in chapter 1.

Subjects and data for the study were drawn from the National Longitudinal Survey of Youth (NLSY). All data were derived from Release 9.0 of the database (NLSY, 1996). This was the most current data set available in the fall of 1998. It included all information collected from inception of the NLSY in 1979 through the 1996 survey year.

The data analyzed were entirely archival in nature. For this reason, the study was exempt from the requirement that investigations involving human subjects receive approval from an institutional review board.

The NLSY Participants

Subjects were drawn from a pool of approximately 6,000 female NLSY participants. The NLSY is sponsored by the U.S. Department of Labor and has been administered since inception by the Center for Human Resource Research (CHRR) at Ohio State University (CHRR, 1995). The National Opinion Research Center (NORC) assists with sample selection, development of survey documents, and interviewing. NORC is affiliated with the University of Chicago.

All NLSY participants were born between 1957 and 1964. They were non-institutionalized and living in the United States at the time of their first interview. The original study population included a total of 12,686 males and females. All fulfilled criteria for one of three subsamples:

* a cross-sectional, civilian sample (*n* = 6,111) selected to be representative of youths living in the United States in 1979;
* a supplemental civilian group (*n* = 5,295) that oversampled for Black, Hispanic, and economically disadvantaged White youth;
* a group of enlisted military personnel (*n* = 1,280) serving in the Army, Navy, Air Force, or Marines as of September 30, 1978.

The process of selecting military participants began with an analysis of personnel records at the U.S. Department of Defense (Bock & Moore, 1986). Personnel born during eligible years were classified by branch of service and by geographic region. Subjects were then selected from each group in a manner providing appropriate representation for each service branch and each geographic area.

Standard area probability methods were used to select nonmilitary participants (Bock & Moore, 1986). In accordance with this technique, progressively smaller areas of residence--counties, census blocks, dwelling units--were sampled randomly to generate a subset of households representative of the nation as a whole. Using these techniques, NORC identified 81,000 households as potential sources of participants.

NORC screeners then contacted each household to assess the willingness of identified families to participate in the survey and to determine the number of eligible youth living in each home. At the conclusion of this process, a total of 11,406 civilian subjects had volunteered to participate. NORC personnel interviewed these subjects at length in 1979. Interviews took place in the range of settings where subjects lived, from private homes and college dormitories to military installations and prisons.

Participants were reinterviewed annually by NORC from 1980 through 1994. Follow-up interviews were sometimes conducted by telephone, but face-to-face interviews in the subject's residence were favored. The study continues at this time. However, since 1994, reassessments have taken place every 2 years instead of annually.

Since entry, participants have provided extensive data on their family structure, attitudes, personal behavior, education, and work experience. NLSY administrators also obtained additional information about participants from their schools and from law enforcement agencies. Data have been coded, reviewed for accuracy, and entered into an archival database.

The sample has changed somewhat over time. No new participants have been added. However, some participants have been lost to follow-up; several hundred have died (CHRR, 1997). Such consequences are common in longitudinal studies of a cohort group.

Downsizings have also caused significant changes to the sample. In 1985, most subjects from the military subsample were terminated. Disadvantaged Whites from the supplemental sample were dropped in 1994. These events left 8,636 of the original 12,686 subjects active as of 1996.

The NLSY Database: Accessibility and Utility

The NLSY data are in the public domain. Interested persons may purchase the data on magnetic tape or CD-ROM. Release 9.0 costs $50 for two CD-ROMs containing the data collected from 1979 through 1996.

A companion set of data, known as the Geocode, includes a broad range of demographic information regarding the Standard Metropolitan Statistical Area (or other population unit) where each subject resides. This file contains information such as crime rates, physician ratios, poverty data, percentage of households receiving governmental assistance, and racial composition. The Geocode is not in the public domain. It is available only to researchers affiliated with institutions who execute a contract with the U.S. Government regarding its use (BLS, 1998).

The NLSY data set has been widely used by researchers in the social sciences. The Center for Human Resource Research maintains an on-line compendium of all publications, presentations, and dissertations based on NLSY data (CHRR, 1998). The compilation is searchable by keyword. As of September 1998, the index included more than 1,500 entries.

Investigations based on the NLSY data span a broad range of subject areas. A review of the on-line bibliography reveals publications that analyzed NLSY data for information about substance abuse, labor force participation, father absence, and reproductive behaviors. Several dozen reports address issues related to vocational aptitudes and cognitive abilities.

Most of the latter analyze results of the Armed Services Vocational Aptitude Battery (ASVAB). The database contains ASVAB scores for almost 12,000 of the original participants. These results, along with extensive life history data on the participants, allowed for testing the research hypotheses posed in this dissertation.

Description of the ASVAB

The Armed Services Vocational Aptitude Battery was developed by the U.S. military services to measure general and specific areas of ability and knowledge. All applicants for enlistment into the armed forces must take the ASVAB to establish eligibility and determine suitability for particular training assignments (CHRR, 1995). The military services have used the test for three decades. It is not commercially available.

In 1980, NLSY subjects were asked to sit for the ASVAB. The U.S. Department of Defense solicited their participation in order to confirm that its test norms for military recruits accurately reflected the abilities of the enlistment-eligible population (Bock & Moore, 1986). Representatives of NORC administered the test to the NLSY subjects at more than 400 sites in the U.S. and overseas. Approximately 94% of the NLSY participants--11,914 subjects--completed the test and received a payment of $50 for their time.

The ASVAB consists of 10 subtests. These components--and the areas covered--are (Kass et al., 1982):

* general science--knowledge of life and physical sciences;
* arithmetic reasoning--ability to solve word problems requiring mathematical reasoning;
* word knowledge--basic vocabulary;
* paragraph comprehension--ability to extract information from written passages;
* numerical operations--very simple arithmetic problems;
* coding speed--ability to match words with six-digit numbers;
* auto and shop information--familiarity with auto mechanics, tools, and shop procedures;
* mathematics knowledge--knowledge of geometry, algebra, and fractions;
* mechanical comprehension--understanding of mechanical principles such as the workings of gears, pulleys, hydraulics as presented in both verbal and pictorial formats;
* electronics information--familiarity with principles of radio and electronics presented verbally or pictorially.

Tables 2 and 3 provide sample questions from each subtest.

In addition to the 10 subtest scores, several composites are calculated. Most important is the Armed Forces Qualifying Test (AFQT) score. The AFQT assesses general ability, as well as potential trainability (Bock & Moore, 1986; Earles & Ree, 1992). The formula for calculating the AFQT includes scores from four subtests: arithmetic reasoning, mathematics knowledge, paragraph comprehension, and word knowledge.

While the AFQT is a good measure of academic ability, other subtests also assess college potential. Mechanical comprehension (MC) predicts ability to succeed in technical areas such as architecture and engineering (DMDC, 1998). General science (GS) is relevant for those interested in careers in the life or physical sciences.

No subtest specifically measures spatial ability. However, mechanical comprehension provides some indicator of spatial competence (Bock & Moore, 1986). Factor analysis of the ASVAB reveals that the mechanical comprehension also reflects quantitative ability, albeit to a lesser extent than arithmetic reasoning and mathematical knowledge (Kass et al., 1982).

Table 2
Sample Questions from ASVAB Subtests Used to Compute AFQT Score

Paragraph Comprehension (PC)	***Arithmetic Reasoning (AR)***
When people move into a new town and become aware of specific features of the area, they soon learn that towns vary as much as people. No two towns are the same; no two states are the same. Hence people must solve living problems in new ways. Also, just as no two towns are the same, no two people have the same problems. According to the passage, towns a. are all the same b. vary only slightly c. vary as much as people do d. are similar in the same state	Two partners, X and Y, agree to divide their profits in the ratio of their investments. If X invested $3000 and Y invested $8000, what will be Y's share of a $22,000 profit? a. $8,250 b. $16,000 c. $6,000 d. $5,864
Mathematics Knowledge (MK)	***Word Knowledge (WK)***
If j and k are positive whole numbers and j+k=12, what is the greatest possible value of jk? a. 6 b. 36 c. 32 d. 11	Solitary most nearly means a. sunny b. being alone c. playing games d. soulful

No effort is made to minimize sex differences in outcome on the ASVAB (Jensen, 1998). This is contrary to the practice of many test developers, who often discard questions with large sex differences in the percentage of persons who answer correctly (Nyborg, 1994). Alternatively, test developers sometimes revise questions if doing so reduces sex-related differences in ability to answer accurately (Willingham & Cole, 1997). Because the ASVAB makes no such adjustments, it may measure cognitive sex differences more accurately than many commercially available tests.

Females outperform males on 4 of the 10 ASVAB subtests. Mean scores are higher for females on the two verbal measures (particularly paragraph comprehension) and on the two speeded tests--numerical operations and coding speed (Bock & Moore, 1986). By contrast, males outperform females on arithmetic reasoning, mathematical knowledge, and subtests of technical knowledge (GS, AS, EI, MC).

Of the 10 subtests, the sex difference is greatest for auto and shop information. This subtest shows an effect size of 1.02 (in favor of males). Sex differences are also large for mechanical comprehension and electronics information. Both have an effect size of .72 (Hedges & Nowell, 1995).

Validity and Reliability of the ASVAB

Like any cognitive test, the ASVAB might be considered a subject of some controversy. Whether any truly informed controversy exists about its value is nonetheless questionable. Since its development, the ASVAB has been tested repeatedly for reliability and predictive validity.

Welsh, Kucinkas, and Curran compiled studies examining the construct, content, and criterion validities of the ASVAB from its first use in 1966 through the late 1980s. This review affirmed its validity for selecting and classifying military candidates (Welsh, Kucinkas, & Curran, 1990). Subsequent investigation also showed that the test accurately predicted scholastic and job performance among military personnel (Earles & Ree, 1992).

As the military has the sole authority to administer the ASVAB, its predictive validity in other situations is moot. Nonetheless, its predictive value relative to high school performance has been investigated. Composite scores derived from the ASVAB proved to be valid predictors of high school grades (Fairbank, Welsh, & Sawin, 1990). However, predictive validity was greater for academic than vocational courses.

Table 4 provides correlations between certain ASVAB subtests and corresponding sections of the Scholastic Aptitude Test. These were generated in conjunction with this research using SAT scores of approximately 900 NLSY subjects. In light of the strong predictive validity of SAT (Willingham & Cole, 1997), these high correlations further demonstrate the validity of the ASVAB as a measure of academic potential.

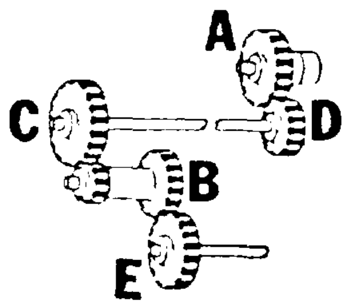
Table 3
Sample Questions from Subtests Not Included in AFQT Score

General Science (GS)	***Numerical Operations (NO)***
We see the moon mainly by a. light reflected from earth b. glow from the moon's phosphere c. starlight d. light from the sun	3x3= a. 6 b. 0 c. 9 d. 1
Electronics Information (EI)	***Auto & Shop Information (AS)***
Which of the following is measured in watts? a. electrical power b. electrical resistance c. electrical potential	Which of the following mechanical devices is NOT a part of the combustion system of an automobile? a. piston b. carburetor c. brake cylinder d. fuel pump

Mechanical Comprehension (MC)

Gear C is on the same shaft as gear D. Two gears that turn in the same direction when the device is turned on are

a. A and C
b. B and C
c. C and E
d. B and E



Coding Speed (CS):

green............	2715	man..............	3451	salt..............	4586
hat................	1413	room............	2864	tree..............	5972

Match the words below to the proper number from the key above.

	A	B	C	D	E
room	1413	2715	2864	3451	4586
green	2715	2864	3451	4586	5972
tree	1413	2715	3451	4586	5972
hat	1413	2715	3451	4586	5972
salt	1413	2864	3451	4586	5972

Reliability of the ASVAB is defined as the ratio of the non-error variance to total score variance in the test population (Bock & Moore, 1986). Re-analysis of reliability values takes place with each new version of the test. Reliability values for the version of the test administered to NLSY participants appear in Table 5 (Bock & Moore, 1986).

Concerns about sex bias are often raised in relation to standardized tests (Jensen, 1998). The ASVAB has been examined thoroughly for evidence of bias against racial minorities and/or females. No such bias has been found (Bock & Mislevy, 1981; Bock & Moore, 1986; Earles & Ree, 1992). To the contrary, the ASVAB--as well as aptitude batteries that include its constituent subtests--predicts that females and minorities will perform better in training than they actually do (Caretta, 1997).

Table 4
ASVAB-SAT Correlations

Tested Correlations ASVAB	SAT	n	r	Sig
Arithmetic Reasoning	Math	920	.806	p < .01
Math Knowledge	Math	920	.772	p < .01
Word Knowledge	Verbal	917	.724	p < .01
Paragraph Comprehension	Verbal	917	.598	p < .01

Table 5
Reliabilities of the ASVAB

Subtest	Reliability
General Science	.86
Arithmetic Reasoning	.87
Word Knowledge	.86
Paragraph Comprehension	.68
Numerical Operations	.71
Coding Speed	.82
Auto and Shop Knowledge	.83
Mathematics Knowledge	.84
Mechanical Comprehension	.83
Electronics Information	.80

Moreover, this investigation was limited to within-sex comparisons. Therefore, concerns about sex bias--whether groundless or legitimate--would have been moot.

Other Test Instruments

In addition to the ASVAB, four psychological rating scales were used in this research. These included the Rosenberg Self-Esteem Scale (RSE), the Center for Epidemiologic Studies-Depression scale (CES-D), the Pearlin Mastery Scale (PMS), and the Scale of Traditional Attitudes (STA). The first three are well-known tests with a long history of use in practice and research. The STA is a little-known scale developed specifically for the NLSY.

Description. All four instruments are self-report inventories. Each provides a list of statements followed by a four-point Likert scale. Respondents circle the choice most applicable to their sentiments. For the RSE, PMS, and STA, the four choices correspond to strongly agree, agree, disagree, or strongly disagree. For the CES-D, which inquires about symptoms during the week previous to testing, the options correspond to none of the time, a little of the time, some of the time, or most of the time.

The number of items varies among the scales. The CES-D has 20; the RSE, 10; and the PMS, 7. The STA actually consists of eight items. However, for reasons explained below, one of the eight questions was excluded from this research.

Tables 6-9 show each of the four scales.

Cost and Accessibility. All four tests are in the public domain and may be used without charge. The Internet provides ready access to the CES-D and RSE (Anonymous, 1998; Measurement-Group, 1997). The Pearlin Mastery Scale can be accessed through the professional literature (Pearlin et al., 1981; Pearlin & Skaff, 1996). Its seven test items also appear in the NLSY database and the 1992 NLSY questionnaire. The latter can be ordered from CHRR at a cost of $13. The Scale of Traditional Attitudes appears in publications available free of charge from the Center for Human Resource Research (CHRR, 1997).

Reliability and Validity. The reliability of the RSE is well-established (Demo, 1985). Internal consistency of test items is high (Hagborg, 1996). In fact, a reliability analysis based on the NLSY sample yielded a Cronbach's alpha of .82 (Plotnick & Butler, 1991).

Validity of the RSE has been demonstrated across a range of predictive outcomes. These include problem behaviors and difficulties with school and/or family (Bagley, Bolitho, & Bertrand, 1997). As a recognized standard for the measurement of self-esteem, the RSE is now used as a benchmark for testing the validity of newer instruments (Lorr & Wunderlich, 1986; Overholser, Schubert, Foliart, & Frost, 1993).

The CES-D scale shows high reliability as well. Analysis of internal consistency in four sample populations yielded reliabilities of .84 to .90 as measured by Cronbach's alpha (Radloff, 1977). Field tests have demonstrated good validity as assessed by correlations with objective criteria for detection of depression. These include clinician ratings, admission for in-patient treatment, or other self-report scales that measure psychological well-being (Radloff, 1977).

The validity of the CES-D has also been established cross-culturally. Strong validity has been reported across samples from a diverse range of populations, including Koreans (Cho & Kim, 1998), citizens of Hong Kong (Cheung & Bagley, 1998), Italians (Fava, 1983), and Mexican-Americans (Roberts, Vernon, & Rhoades, 1989).

Table 6
The Rosenberg Self-Esteem Scale (RSE)

ITEM #	STATEMENT	STRONGLY AGREE 1	AGREE 2	DIS-AGREE 3	STRONGLY DISAGREE 4
1.	I feel that I'm a person of worth, at least on an equal plane with others.	1	2	3	4
2.	I feel that I have a number of good qualities.	1	2	3	4
3.	All in all, I am inclined to feel that I am a failure.	1	2	3	4
4.	I am able to do things as well as most other people.	1	2	3	4
5.	I feel I do not have much to be proud of.	1	2	3	4
6.	I take a positive attitude toward myself.	1	2	3	4
7.	On the whole I am satisfied with myself.	1	2	3	4
8.	I wish I could have more respect for myself.	1	2	3	4
9.	I certainly feel useless at times.	1	2	3	4
10.	At times I think I am no good at all.	1	2	3	4

Table 7
The Pearlin Mastery Scale

ITEM #	STATEMENT	STRONGLY DISAGREE 1	DIS-AGREE 2	AGREE 3	STRONGLY AGREE 4
1.	There is really no way I can solve some of the problems I have.	1	2	3	4
2.	Sometimes I feel that I'm being pushed around in life.	1	2	3	4
3.	I have little control over the things that happen to me.	1	2	3	4
4.	I can do just about anything I really set my mind to.	1	2	3	4
5.	I often feel helpless in dealing with the problems of life.	1	2	3	4
6.	What happens to me in the future mostly depends on me.	1	2	3	4
7.	There is little I can do to change many of the important things in my life.	1	2	3	4

The PMS was tested on a mixed-sex sample of 2,300 people representative of the Chicago area population during the 1970s. Inter-item correlations were high, particularly among those statements that are direct expressions of helplessness over the forces controlling one's life (Pearlin et al., 1981). Re-administration of the scale to the same individuals after a substantial period of time yielded moderately consistent results (r = .44).

To test internal consistency of the PMS, a reliability analysis was performed in conjunction with this research. Responses of 8,938 NLSY participants were used. Results yielded a Cronbach's alpha of .79.

The NLSY database also contains results from an abbreviated version of the Rotter Locus of Control Scale. The questions on this shortened scale also inquire about the respondent's sense of personal control over life events. The four-item scale has Cronbach's alpha of only .29 (Plotnick & Butler, 1991). The PMS therefore offers a far more reliable measure of internal control.

The developers of the PMS did not attempt to establish predictive validity. However, it has since been found to predict compliance with mental health treatment (Bowen, South, Fischer, & Looman, 1994). In an effort to streamline evaluation of patients with high-risk pregnancies, Goldenberg pooled the 59 items of the PMS, RES, CES-

D, and two additional scales for measuring stress and anxiety. Factor analysis was used to eliminate redundancy and establish primary factors. The result was an abbreviated scale of 28 items that predicted birth weight, gestational age, fetal growth restriction, and pre-term delivery as well as the full pool of 59 items (Goldenberg et al., 1997).

NLSY statisticians have not validated the STA. Its internal consistency has been tested twice. One analysis found strong correlations between five of the eight items, but little correlation between these five and the remaining three (Shapiro & Crowley, 1981). A second analysis reported a Cronbach's alpha of .68 for a modified version of the scale that used seven of its eight questions. As the latter version has been used in several research studies, it was also adopted here.

Scoring. All four tests are readily hand-scored by the administrator. Scoring is largely a matter of summing the responses. However, researchers must adjust values for reverse-scored items prior to computing the total score. Alternatively, one can use a formula that takes reverse-scored items into account. For purposes of this research, formulas making necessary adjustments were adopted. These can be found in this chapter under Data Analysis.

Table 8
The CES-D Depression Scale

ITEM #	Instructions: Read each statement, then circle the number that best describes how often you felt that way during the past week.	Rarely; none of the time; 1 day 0	Some; a little of the time; 1-2 days 1	Occasionally; moderate amount of the time; 3-4 days 2	Most/all of the time; 5-7 days 3
1.	I was bothered by things that usually don't bother me.	0	1	2	3
2	I did not feel like eating; my appetite was poor.	0	1	2	3
3	I felt that I couldn't shake off the blues even with help from my family and friends.	0	1	2	3
4.	I felt that I was just as good as other people.	0	1	2	3
5	I had trouble keeping my mind on what I was doing.	0	1	2	3
6.	I felt depressed.	0	1	2	3
7.	I felt that everything I did was an effort.	0	1	2	3
8.	I felt hopeful about the future.	0	1	2	3
9.	I thought my life had been a failure.	0	1	2	3
10.	I felt fearful.	0	1	2	3
11.	My sleep was restless.	0	1	2	3
12.	I was happy.	0	1	2	3
13.	I talked less than usual.	0	1	2	3
14.	I felt lonely.	0	1	2	3
15.	People were unfriendly.	0	1	2	3
16.	I enjoyed life.	0	1	2	3
17.	I had crying spells.	0	1	2	3
18.	I felt sad.	0	1	2	3
19.	I felt that people dislike me.	0	1	2	3
20.	I could not get going.	0	1	2	3

Table 9
The Scale of Traditional Attitudes

ITEM #	STATEMENT	STRONGLY DISAGREE 1	DIS-AGREE 2	AGREE 3	STRONGLY AGREE 4
1.	A woman's place is in the home, not in the office or shop.	1	2	3	4
2.	A wife who carries out her full family responsibilities doesn't have time for outside employment.	1	2	3	4
3.	A working wife feels more useful than one who doesn't hold a job.	1	2	3	4
4.	Employment of wives leads to more juvenile delinquency.	1	2	3	4
5.	It is much better for everyone concerned if the man is the achiever outside the home and the woman takes care of the home and family.	1	2	3	4
6.	Men should share the work around the house with women, such as doing dishes, cleaning, and so forth.	1	2	3	4
7.	Women are much happier if they stay at home and take care of their children.	1	2	3	4

Norms and interpretation of scores on these instruments are left to the discretion of the test administrator. However, the developers of the CES-D indicate that a score of 16 or higher suggests a significant degree of depression (Radloff, 1977). This norm had no relevance to this research, as scales were used for comparative rather than diagnostic purposes.

Utility. Researchers and clinicians have used the CES-D extensively since its release in 1977. It is employed in evaluation and follow-up of HIV-positive patients (Measurement-Group, 1997); assessment of symptom severity over time in psychiatric patients (Weissman, Sholomskas, Pottenger, Prusoff, & Locke, 1977); and screening for depression among patients with disruptions of mood secondary to cerebrovascular events (Weissman et al., 1977). As would be expected, it is also used to measure the incidence of depression in general populations (Schulberg et al., 1985).

Uses of the RSE have similarly been varied and extensive. A common use is simply to assess self-esteem. The scale has shown utility for this purpose in a wide range of populations--from adolescent mothers (Barth, Schinke, & Maxwell, 1983) to heavy drinkers (Ratliff & Burkhart, 1984; Steffenhagen & Steffenhagen, 1985). Similarly, investigators have used the scale to measure both between-sex and within-sex differences in self-esteem (Lamke, 1982). Notably, females with less stereotypically feminine characteristics had higher self-esteem as assessed by the scale.

The RSE also shows predictive value for assessing willingness to choose and persist toward difficult academic goals (Leondari, Syngollitou, & Kiosseoglou, 1998; Levy & Baumgardner, 1991) as well as likelihood of compliance with self-care recommendations (Conn, Taylor, & Hayes, 1992). In the latter example, adherence to recommended diet, exercise, and stress reduction measures correlated positively with self-esteem as measured by the RSE. Compliance with anti-smoking advice, however, was unrelated to RSE score.

The PMS has been used primarily to gauge the effects of stress on sense of personal control and to assess coping ability (Pearlin et al., 1981; Pearlin & Skaff, 1996). The scale has been used repeatedly in studies of caregiver stress (Aneshensel, Pearlin, & Schuler, 1993; Pearlin, Aneshensel, & LeBlanc, 1997; Skaff & Pearlin, 1992; Skaff, Pearlin, & Mullan, 1996). It has also been used to study how varying types of stress affect sense of internal control (Pearlin et al., 1981). Economic strain, for instance, shows very strong effects on this aspect of well-being. The scores of those who are employed and/or have higher household incomes tend to show a greater sense of internal control (Walford-Kraemer & Light, 1984).

The STA has had comparably little use. It has been employed to assess differences in sex-role attitudes among various demographic groups in the United States. One such analysis found that males had more traditional attitudes about the role of females than did women themselves (Shapiro & Crowley, 1981). In addition, both Blacks and Whites expressed less traditional attitudes about the role of women than did Hispanics.

Other researchers have used the scale to determine whether sex-role attitudes correlate with the participation of Hispanic females in the labor force; results suggested education to be the more salient factor (Ortiz & Cooney, 1984). Investigators have also used the scale to assess risk of adolescent pregnancy. Females whose responses suggested strongly traditional attitudes were more likely to become unwed mothers (Plotnick & Butler, 1991).

All other variables used in this research were derived from responses to individual questions rather than rating scales.

Subject Selection and Retrieval of Data

Using software included within the NLSY database, all variables necessary to analyze the questions posed in chapter 1 were extracted from the CD-ROM. The extract was written to a separate database in the form of a dBase file, then transferred into SPSS using its import function. All analyses thereafter were executed with the SPSS software package (release 8.0 for Microsoft Windows).

A frequency analysis was carried out on the ASVAB subtest scores for mathematical knowledge, arithmetic reasoning, and mechanical comprehension. These analyses established decile ranges for the three subtests. Female subjects scoring in the top decile on all three tests were tagged for inclusion in the HMHM group.

The control group consisted of females whose ASVAB scores met the following criteria:

* an AFQT score greater than or equal to the 75th percentile;
* a mathematical knowledge score below the top decile;
* an arithmetic reasoning score below the top decile;
* a mechanical reasoning score less than or equal to the 55th percentile for this subtest.

Mean ages for each ability group were calculated to insure similarity between the groups. Demographic data such as socioeconomic status and ethnicity were also compiled, as were scores on a variety of ASVAB subtests not included in the selection criteria.

A note about these criteria is in order. The study was limited to academically able individuals for several reasons. The first was a desire to study subjects representative of the college-capable population. Women who can succeed in technical professions are unlikely to be found among persons of lesser general ability.

Also relevant was evidence that cognitive patterns are more variable at higher levels of ability (Deary & Pagliari, 1991; Detterman, 1991; Detterman & Daniel, 1989). Persons of low ability have a greater predilection to perform similarly across widely disparate cognitive tests. Identifying persons with unusual cognitive patterns is consequently more difficult among such populations.

Several investigators have confirmed this phenomenon using the same subjects and aptitude battery employed in this research (Evans, 1998; Legree, Pifer, & Grafton, 1996). In both analyses, persons with higher levels of general ability showed more variable cognitive patterns. Accordingly, this sector provided the best opportunity to isolate the ability pattern in question and examine its correlates.

Data Analysis

Chapter 1 set forth a series of descriptive statements regarding the relationship between the HMHM ability pattern and various aspects of life history. Analysis for each was as follows.

1. Height will be greater in HMHM females than in controls. Using data on adult height (in inches) provided by respondents in 1985, the means and standard deviations were calculated for each ability group. A one-tailed *t* test was used to determine whether differences between the two groups were statistically significant.

2. Body fat will be lower in HMHM females than in controls. As explained in chapter 1, body fat was operationalized as the Body Mass Index (BMI). This variable was calculated by dividing a subject's weight in pounds by the square of her height in inches and multiplying the result by 703.1. On the assumption that virtually all subjects had reached their full adult height as of 1981, a BMI value for that year was calculated using height data from 1985 and weight data for 1981.

The BMI was also computed for 1986, 1990, and 1996. Means and standard deviations for each ability group were computed and analyzed by one-tailed *t* test.

3. Maturational tempo will be slower in HMHM females than controls pattern. Age at menarche was the first marker for maturational tempo. Means and standard deviations for each ability group were computed for recalled age at menarche and analyzed by one-tailed *t* test. The second marker of maturational tempo, age at first intercourse, was similarly analyzed by computing means and standard deviations for each ability group. Results were evaluated by one-tailed *t* test.

4. Physical activity will be greater in HMHM females than controls. Activity was operationalized as participation in high school sports or cheerleading. In the NLSY database, these responses are in the form of a two-category variable; subjects simply answered yes or no when asked in 1984 about participation in various extracurricular activities. Responses were tabulated for the two ability groups. A one-sided Fisher's exact test was used to analyze the results.

5. Traditional sex-role attitude will be less pronounced among HMHM females than controls. Using the convention t1 through t7 to represent the Likert value chosen on the individual items of the STA, total scores for each subject were computed using the formula:

10 + t1 + t2 - t3 + t4 + t5 - t6 + t7.

Means and standard deviations for total STA scores were calculated for each ability group and analyzed by one-tailed *t* test.

Group means and standard deviations were also calculated for number of children wanted and age at which subjects expected to marry. These were analyzed by two-tailed *t* test because no specific prediction of directionality was made for either variable. Several two-category variables were also analyzed: desired lifestyle at age 35 (working vs. married with family), desire to marry (yes vs. no), and desire to remain childless (yes vs. no). The resulting 2 x 2 tables were analyzed by chi-square test.

Occupational codes for the years 1985, 1990, and 1996 were translated into job titles using information provided in the NLSY Codebook Supplement (CHRR, 1996). All employment fields were then recoded as female-typical, sex-neutral, or male-typical.

Occupations classified as female-typical included clerical positions, allied health professions, elementary and secondary school teaching, childcare, hair styling, and food service occupations. Administrative, managerial, and sales positions were categorized as sex-neutral. Jobs in engineering, construction, heavy industry, and computer programming were coded as male-typical. Statistics on the sex of individuals employed in specific fields were consulted as necessary to determine appropriate category for reclassification of subjects' occupations (Wooten, 1997). Data were analyzed by chi-square test.

6. Parity will be lower among HMHM females than among controls. For both ability groups, means and standard deviations were calculated for number of live births as of 1984, 1988, and 1996. Results were analyzed by one-tailed *t* test. A new two-category variable--childless as of 1996--was created from the existing data on number of births. Yes-no responses were organized into a 2 x 2 table and analyzed by one-sided Fisher's exact test.

7. Number of pregnancies will be lower among HMHM females than controls. As noted in chapter 1, the number of pregnancies reported by participants included not only those resulting in live births, but also those terminated by spontaneous or induced abortion. Means and standard deviations for this variable were calculated for the two ability groups for 3 survey years and evaluated by one-tailed *t* test. A new two-category variable--ever pregnant as of 1996--was created from the existing data on number of births. Yes-no responses were organized into a 2 x 2 table and analyzed by one-sided Fisher's exact test.

8. Libido in adulthood will be higher among HMHM females than among controls. Means and standard deviations for monthly coital frequencies reported in 1984 and 1985 were computed for the two ability groups and analyzed by one-tailed *t* test.

9. Religiosity will be lower among HMHM females than among controls. Subjects reported their church attendance on an ascending numeric scale; 1 represented little or no attendance, while 6 indicated very frequent attendance. Means and standard deviations for scores on this scale were computed for the two ability groups and analyzed by one-tailed *t* test.

10. Locus of control will be more internal among HMHM females than controls. Using the convention p1 through p7 to represent values circled by participants on each of the questions on the PMS, total scores were computed as:

25 - p1 - p2 - p3 + p4 - p5 + p6 - p7.

(Under this formula, progressively higher scores indicated increasing levels of internality.) Mean total scores and standard deviations were calculated for each ability group and evaluated by one-tailed *t* test.

11. Left-handedness will be more common among HMHM females than among controls. The number of right- and left-handed subjects in each ability group was organized into a 2 x 2 table and analyzed by a one-sided Fisher's exact test.

12. Depressive symptoms will be less common among HMHM females than controls. Using the convention d1 through d20 to represent the individual item responses, subjects' total scores on the CES-D were computed using the formula:

12 + d1 + d2 + d3 - d4 + d5 + d6 +d7 - d8 + d9 + d10 + d11 - d12 + d13 + d14 + d15 - d16 + d17 + d18 + d19 + d20.

Calculated this way, rising total scores indicate increasing severity of depressive symptoms. Means and standard deviations for the CES-D score were computed for each ability group and assessed by one-tailed *t* test.

13. Self-esteem will be higher among HMHM females than among controls. Using the convention that s1 through s10 correspond to the value circled by respondents on the RSE, total scores for each subject were calculated using the formula:

25 - s1 - s2 + s3 - s4 + s5 - s6 - s7 + s8 + s9 + s10.

Means and standard deviations were calculated for both groups and analyzed one-tailed *t* test.

Table 10 summarizes scoring information for the Rosenberg scale and the other three psychological inventories used in this research.

In some cases, outcomes suggested additional avenues of inquiry. These were investigated as data allowed and are discussed in the next chapter along with the findings from the analyses described above.

As discussed in chapter 1, this inquiry was less concerned with specific variables than with the validity of the overall model. In this context, and due to progressive loss of statistical power resulting from downsizing of the data-base, an alpha value of .10 or less was used as the standard for statistical significance.

Table 10
Score Ranges for Psychological Inventories

Scale	Minimum Score	Maximum Score	Increasing Score Indicates
Rosenberg Self-Esteem Scale (RSE)	10	40	Higher self-esteem
Pearlin Mastery (PMS)	7	28	Higher sense of internal control
CES-D Depression Scale	0	60	Increasing severity of depressive symptoms
Scale of Traditional Attitudes	7	28	Increasingly traditional attitudes about women's roles

Chapter 4: Findings

General Characteristics of Study Groups

Age. Selection criteria generated 67 HMHM females and 60 controls of similar mean age. The mean month and year of birth was May 1959 for HMHM females and February 1959 for controls. While the birth years of subjects ranged from 1957 to 1964, majorities of both groups were born prior to 1960.

Incidence of HMHM pattern by sex. Frequency analysis confirmed the rarity of the HMHM ability pattern in females. Only 1.1% of the 5,927 NLSY females who completed the ASVAB scored in the top decile on each of the three subtests that define the HMHM cognitive pattern. Among males, the corresponding figure was 5.5%. Accordingly, the HMHM ability pattern was five times as common in males as in females.

Distribution of specific abilities. Analysis also revealed that virtually all differences in scoring patterns between HMHM and control females resulted from the better mathematical and technical competence of the former (Table 11). HMHM females outperformed controls by wide margins in quantitative and technical areas predictive of success in scientific fields. Mean scores in the verbal domains of paragraph comprehension and word knowledge did not differ significantly between groups. Similarly, means on the ASVAB verbal composite showed no significant difference between the groups. As a consequence, the two groups were essentially equivalent relative to both age and verbal ability.

Evaluation of scoring patterns suggested that relative strengths of the HMHM group were more similar to those of males than females. A ratio of mathematical to verbal ability was calculated by dividing the arithmetic reasoning score by the paragraph comprehension score. These subtests most closely resemble mathematical and verbal reasoning tasks found on college admissions tests. Among males, the mean ratio was 1.08, compared to .99 for females ($p < .001$; two-tailed t test). For HMHM females, the mean ratio was 1.10, compared to .98 for controls ($p < .001$; two-tailed t test). HMHM males had a mean ratio of 1.12. Thus, the mean for HMHM females fell slightly below that of HMHM males, but exceeded the average for all males.

These ratios should not be confused with total scores. While the relative proportion of mathematical to verbal ability in HMHM females was slightly greater than the mean for all males, the magnitude of these abilities differed greatly. HMHM females are an elite group. Their mean ability scores for all areas shown in Table 11 far exceed averages for either sex.

Table 11
Selected ASVAB Subtest Scores for HMHM and Control Groups

Subtest	HMHM (n = 67)	Control (*n* = 60)	*p* value[a]	Tails[b]
Arithmetic reasoning				
M	64.9	57.1	<.001	1
SD	1.2	2.8		
General science				
M	60.7	54.5	<.001	2
SD	4.6	4.8		
Mathematics knowledge				
M	65.9	58.2	<.001	1
SD	1.6	2.9		
Mechanical comprehension				
M	64.9	43.4	<.001	1
SD	2.1	3.0		
Paragraph comprehension				
M	58.8	58.7	.88	2
SD	2.4	3.2		
Word knowledge				
M	59.6	59.1	.14	2
SD	2.2	1.9		
Verbal composite				
M	59.7	59.2	.17	2
SD	2.0	1.9		

[a] By t test; [b] one-tailed test used for subtests defining eligibility for HMHM group.

Ethnicity. All but one of the 67 HMHM females described their ancestry as European. The lone exception was biracial, with one Black and one Japanese parent. About three-fourths of controls were Caucasian (47 of 60). Of the remaining 13, 7 were Black and 6 Hispanic. Among the latter, 4 listed their ethnicity as Mexican. The remaining 2 reported Cuban ancestry.

Socioeconomic status. Both groups contained a considerable number of subjects from homes specifically classified as poor. One-fourth of HMHM subjects and one-third of controls came from low-SES families. The proportion of individuals from economically disadvantaged homes may have actually been higher because SES of subjects in the military subsample was not provided. Sixteen per cent of HMHM females were from the military subsample, as were 8% of controls.

In stark contrast to stereotypes, about 15% of the HMHM females from economically disadvantaged homes scored at or above the 97th percentile on the AFQT. The AFQT is a good measure of IQ (Jensen, 1998). The relevance of general intelligence to the study results will be explored later in this chapter.

No notable differences in characteristics such as birth order or family size were found. Mean number of siblings was 2.9 for HMHM subjects and 3.0 for controls. About 24% of HMHM and 30% of control subjects were first-borns. Four of 67 HMHM females, and one of 60 controls, were only-children. None of these differences approached statistical significance.

Somatic Characteristics

Height. The mean height of HMHM females exceeded that of controls by almost an inch (Table 12). The difference was statistically significant. Therefore, the prediction that HMHM females would be taller than controls was supported.

Coding heights as short, average, or tall permitted a more graphic illustration of the difference between the two groups. Figure 3 displays the distribution, showing HMHM females less likely than controls to be short, and much more likely to be tall.

Body Mass Index. The model also predicted that HMHM females would be leaner than controls. Despite their greater height, HMHM females tended to weigh less than controls. As a consequence, HMHM females repeatedly showed lower means on the body mass index (Table 12).

In 1981, the difference in mean BMI between groups trended in the direction predicted but fell short of statistical significance. In all three follow-up years, differences between groups were statistically significant. The overall pattern of results therefore affirmed the model's prediction.

BMI increased in both groups with age; however, the magnitude of increase was greater for controls. Of course, changes in weight rather than height accounted for these increases. To consider the possible influence of factors related to childbearing, weight changes for the study periods of 1982-1992 and 1986-1996 were analyzed using parity as a co-variate. A significant interaction between number of children ever born and degree of weight change between 1982 and 1992 was observed, $F(1,72) = 4.6$, $p = .04$. The interaction between parity and weight change was also significant for the 1986-1996 study period, $F(1, 65) = 5.6$, $p = .02$.

Physical activity. HMHM females were much more likely than controls to have participated in athletics or cheerleading during their high school years (Table 13). These differences in participation were statistically significant, affirming the third prediction of the dissertation model.

Table 12
Height and Body Mass Index (BMI) for HMHM and Control Groups

Measurement	HMHM	Control	One-tailed *p* value[a]
Height, 1985 (inches)			
M	65.5	64.7	.06
SD	2.3	2.4	
n	54	52	
BMI, 1981[b]			
M	21.6	22.0	.29
SD	3.7	4.2	
n	54	52	
BMI, 1986			
M	22.5	24.0	.07
SD	4.3	5.6	
n	52	50	
BMI, 1990			
M	23.4	25.2	.06
SD	4.6	6.3	
n	52	45	
BMI, 1996			
M	24.9	27.2	.09
SD	6.0	8.0	
n	35	35	

[a] By *t* test; [b] calculations based on height data reported in 1985; these are presumed, but not guaranteed to be accurate measures of 1981 height.

Handedness. Left-handedness was more common among HMHM females than controls (Table 13). This trend was in the direction predicted but fell short of statistical significance ($p = .22$). However, the very small number of left-handed subjects must be borne in mind.

In the full NLSY sample, 89% of females and 85% of males expressed right-hand preference; $\chi^2(3, N = 9003) = 57.5$, $p < .001$. The percentage for HMHM subjects was intermediate to these values.

Coupled with the results reported in Table 12, these outcomes supported three of the four predictions made regarding somatic traits of HMHM females. Given the small number of left-handed subjects and a *p* value that trended toward, but did not reach statistical significance, findings relative to the prediction about handedness findings were considered inconclusive.

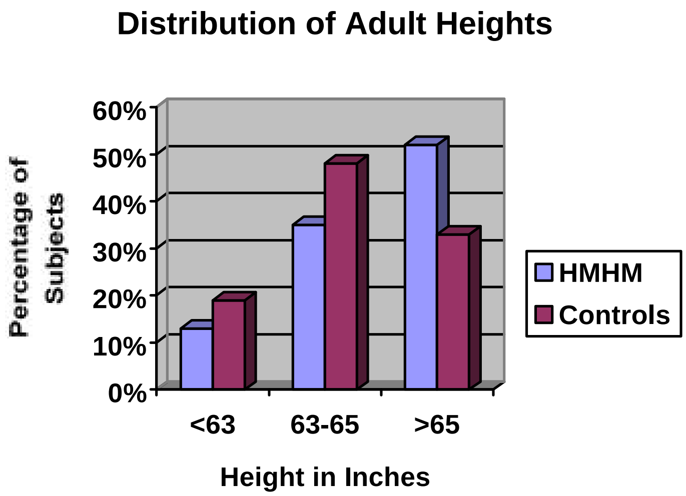


Figure 3. Distribution of adult heights.

Percentage of HMHM females (*n* = 54) and controls (*n* = 52) classified as short, average, or tall.

Table 13
Physical Activity and Handedness for HMHM and Control Groups

Characteristic	HMHM	Control	One-sided *p* value[a]
Participated in high school sports, pep club, or cheerleading			
Yes	62%	44%	.04
No	38%	56%	
n	65	58	
Handedness			
Right	88%	95%	.22
Left	13%[b]	5%	
n	40	40	

[a] By Fisher's exact test; [b] percentages exceed 100 due to rounding.

Psychological Characteristics

Emotional well-being. The dissertation model predicted that HMHM females would have a more internal locus of control, higher self-esteem, and fewer depressive symptoms than controls. Despite a nonsignificant trend suggesting higher internality in the HMHM group, the data did not support these predictions.

Mean scores on the Rosenberg Self-Esteem survey were virtually identical among the groups. The RSE was administered twice. In 1980, controls outscored HMHM females by 0.2 on a 40-point scale. Seven years later, HMHM subjects outscored controls by one-tenth of a point--a nonsignificant result.

Mean score on the CES-D depression scale differed by less than a half point between groups. The trend followed the direction predicted, but was neither statistically significant nor clinically meaningful. Particularly striking were the remarkably low scores of both groups. Mean scores for both HMHM females and controls were about 50% below the level suggested to signal psychological distress.

As noted above, females in the HMHM group outscored controls on the Pearlin Mastery Scale--the measure of internality. The magnitude of the difference was small--a half point on a scale of 28. Although not statistically significant, the direction of the scores corresponded to the dissertation model. Sample size may have been a factor contributing to the lack of statistical significance.

Whether viewed individually or in aggregate, results on these standard measures of psychological status suggest equivalent levels of emotional well-being among the two ability groups. Table 14 summarizes outcomes from these self-report inventories.

Religiosity. Analysis of scores on the NLSY church attendance scale supported the prediction that HMHM females would show less religiosity than controls (Table 15). Subjects were young adults at the time, ranging in age from 17-25. The mean score of 3.0 among controls corresponds to attendance at religious services once a month.

Traditionality. Subjects completed the NLSY Scale of Traditional Attitudes in both 1979 and 1987. Contrary to the model predictions, HMHM females initially scored higher on the traditionality scale than did controls. However, the difference in mean scores among groups was small--0.3 points on a scale of 28 (Table 15).

Table 14
Psychological Inventory Scores for HMHM and Control Groups

Test Instrument	HMHM	Control	One-tailed *p* value[a]
CES-D Depression Scale			
M	7.7	8.1	.41
SD	7.6	8.5	
n	40	39	
Pearlin Mastery Scale			
M	23.2	22.7	.25
SD	3.6	3.3	
n	39	39	
Rosenberg Self-Esteem Scale-1980			
M	34.5	34.7	N/A[b]
SD	3.7	3.9	
n	66	59	
Rosenberg Self-Esteem Scale-1987			
M	35.1	35.0	.46
SD	3.9	3.6	
n	51	51	

[a] By *t* test; [b] direction of results contrary to prediction.

An opposite pattern occurred in 1987, with controls attaining higher mean scores than HMHM females. This was consistent with the dissertation model, but fell short of statistical significance. Moreover, the degree of difference between groups was again small; controls scored only a half point higher than HMHM females. This insignificant difference, coupled with the 1979 findings, did not support the notion that the groups differed in traditionality as measured by the NLSY rating scale. Furthermore, two additional measures of traditional attitudes--number of children wanted and desired age at marriage--were virtually identical for both groups (Table 16).

The percentage of respondents expressing a desire to remain childless was somewhat higher among HMHM females than controls. Similarly, when asked whether she would prefer to be working at age 35 or married with family, the HMHM female was slightly more likely to choose the former option. While suggesting the HMHM group to be somewhat more career-oriented, these results did not approach statistical significance. Moreover, interpreting these responses is difficult, as females can simultaneously be married with family and in the workforce.

Two outcomes did suggest lesser traditionality among HMHM females: marital status and occupation. As of 1996, 28% of HMHM females had never been married, compared to 15% of controls. This difference between the groups fell just short of significance (Table 15). Whether it indicates lesser traditionality among HMHM females is difficult to say; contentedness with marital status was not assessed. Nonetheless, the difference between expressed desire for marriage and actual outcome is notable. On entering the NLSY in 1979, fewer than 4% of HMHM females wished to remain single throughout life.

Table 15
Measures of Religiosity, Traditionality, and Marital History for HMHM and Control Groups

Characteristic	HMHM	Control	One-tailed *p* value
Religious attendance scale			
M	2.6	3.0	.09[a]
SD	1.6	1.7	
n	66	58	
Traditional values scale-1979			
M	13.2	12.9	N/A[b]
SD	3.0	2.9	
n	63	55	
Traditional values scale-1987			
M	12.0	12.5	.25[a]
SD	2.7	3.7	
n	46	42	
Is or has been married-1996			
Yes	72%	85%	.12[c]
No	28%	15%	
n	39	40	

[a] By *t* test; [b] direction of results contrary to prediction; [c] by Fisher's exact test.

The occupational data were compelling. HMHM females were significantly more likely than controls to work in sex-neutral or male-dominated occupations such as computer sciences and engineering. Table 17 shows the distribution of employment for the two groups. Interestingly, the trend away from classically female jobs increased over time for HMHM females. The distribution of job types among controls showed relatively little change.

Viewed as a whole, the traditionality data are too inconsistent to support or reject the original prediction. Therefore, of the five psychological traits in the dissertation model, only the prediction regarding religiosity was supported.

Table 16
Work and Family Aspirations for HMHM and Control Groups

Characteristic	HMHM	Control	*p* value
Number of children desired			
M	2.2	2.3	.72[a]
SD	1.3	1.3	
n	66	60	
Age would like to marry (1-4 scale)[b]			
M	2.5	2.5	.66[a]
SD	0.6	0.7	
n	53	47	
Ideal situation at 35			
Working at some occupation	86%	79%	.35[c]
Married with family	15%[d]	21%	
n	62	57	
Want to remain single for life			
Yes	4%	4%	1.00[c]
No	96%	96%	
n	55	49	
Want no children			
Yes	18%	13%	.45[c]
No	82%	87%	
n	66	60	

[a] By two-tailed *t* test; [b] 1 = <20; 2 = 20-24; 3 = 25-29; 4 = 30 or older; [c] by two-sided chi-square test; [d] percentages exceed 100 due to rounding.

Reproductive Characteristics

Maturational tempo. The dissertation model predicted that HMHM females would mature more slowly than controls. The operational definition focused on two variables: age at menarche and age at first intercourse.

Analysis revealed that control females reached menarche an average 5 months earlier than HMHM females. Controls were also younger at first intercourse. The mean difference between the groups was about 10 months. For both variables, the differences were statistically significant (Table 18).

The mean interval from menarche to sexual debut was 6 months longer in HMHM females than controls. This difference fell short of statistical significance ($p = .17$; one-tailed *t* test). Coupled with other data, it nonetheless reinforces the finding that HMHM females reached reproductive milestones more slowly than controls.

Table 17
Traditionality of Employment for HMHM and Control Groups

Characteristic	HMHM	Control	*p* value[a]
Occupation in 1985			
Female-typical	51%	55%	
Sex-neutral	26%	40%	.02
Male-typical	23%	4%	
n	53	47	
Occupation in 1990			
Female-typical	44%	50%	
Sex-neutral	38%	48%	.06
Male-typical	18%	2%	
n	45	42	
Occupation in 1996			
Female-typical	23%	55%	
Sex-neutral	50%	42%	.006
Male-typical	27%	3%	
n	30	33	

[a] By two-sided chi-square test.

Biologists have long viewed age at reproduction as another marker of maturational tempo (Wilson, 1980). No prediction was made for this variable owing to concern that career paths--such as entering the military or attending college--might confound the results.

Recognizing this possibility, ages at first pregnancy and at first birth were computed for both groups. Controls did report a significantly earlier mean age at first pregnancy. However, some of these pregnancies were terminated. As a consequence, age at first birth was similar between groups.

Number of pregnancies. The dissertation model anticipated fewer pregnancies among HMHM females than controls. Table 19 provides relevant data for 3 survey years: 1984, 1988, and 1996.

Given the relatively young age of some subjects in 1984, data for that year may seem of little value. However, 1984 was the final year of participation for most members of the military subsample. About 15% of HMHM subjects were among that group. Accordingly, the 1984 data provided the best opportunity to include some information, however limited, regarding the reproductive histories of these women.

As of 1984, HMHM females reported fewer pregnancies than controls. Differences fell just short of statistical significance. By 1988, a significant difference emerged between the two groups and persisted through the 1996 survey year (Table 19). These cumulative results logically carry more weight than the 1984 data. The prediction that HMHM females would experience fewer pregnancies was therefore supported. In fact, this proved to be one of the more dramatic differences between the groups.

Table 18
Markers of Maturational Tempo for HMHM and Control Groups

Age at	HMHM	Control	*p* value
menarche			
M	12.8	12.4	.05[a]
SD	1.3	1.3	
n	65	58	
1st intercourse			
M	18.9	18.1	.05[a]
SD	2.4	2.4	
n	52	50	
1st pregnancy			
M	27.2	25.0	.09[b]
SD	3.7	5.1	
n	18	33	
birth of 1st child			
M	27.2	27.0	.87[b]
SD	4.0	4.6	
n	20	31	

[a] By one-tailed *t* test; [b] by two-tailed *t* test.

Contributing to the notable difference in pregnancy rates between the groups was the large percentage of HMHM females who had never conceived at all. As of 1996, 45% of HMHM females had never been pregnant, compared to 15% of controls. Of all pregnancy-related outcomes, this one was particularly unlikely to have occurred by chance (Table 19).

As of 1996, the mean age for both groups was 37. Most subjects therefore remain capable of pregnancy, yet a large number of them are past their most fertile years. In this context, HMHM females appear unlikely to close the pregnancy gap between themselves and controls.

Contraceptive use was examined as part of an effort to account for the pregnancy differential. Although extensive data are not available, interviewers regularly inquired about use of contraception by subjects and/or their partners. Participants were asked to complete a form that asked, did you or your partner use birth control during the past month? A no answer essentially affirmed sexual activity. Abstinent individuals were coded separately from respondents who indicated non-use of contraception.

Table 19
Longitudinal Pregnancy Histories for HMHM and Control Groups

Characteristic	HMHM	Control	1-tailed *p* value
# pregnancies by 1984			
M	0.6	0.8	.13
SD	1.0	1.1	
n	65	57	
# pregnancies by 1988			
M	0.8	1.6	.008
SD	1.3	1.8	
n	51	50	
# pregnancies by 1996			
M	1.5	2.4	.02
SD	1.8	1.8	
n	38	38	
Ever pregnant by 1996			
Yes	55%	85%	.005[a]
No	45%	15%	
n	38	39	

[a] By one-sided Fisher's exact test.

A cumulative approach was adopted to evaluate these data. A contraceptive use composite score was created using responses for the even-numbered years beginning with 1984 and ending with 1996. (Data from 1985 were excluded due to a large amount of missing information.) Each yes answer was assigned a value of one; a participant who consistently used contraception in the month prior to interview had a score of 7. Subjects pregnant during any of the time periods in question were excluded from the analysis.

Results showed no significant difference in overall use of contraception between the groups (Table 20). Accordingly, the substantial difference in pregnancy rates between HMHM females and controls could not be attributed to the former using contraception more frequently. Mean scores indicated somewhat less use of contraceptives by HMHM females. However, differences were not statistically significant.

To examine whether the fertility differential might have resulted from HMHM females not attempting pregnancy until their older, less fertile years, two subscores for contraceptive use were calculated. The first represented usage during the 1984, 1986, and 1988 study years. The second spanned the even-numbered years of the 1990s. Again, no significant differences were found between the groups (Table 20). Any trend, in fact, favored lesser use of contraception by HMHM females.

Table 20
Sexual Frequency and Contraceptive Use for HMHM and Control Groups

Characteristic	HMHM	Control	*p* value
Monthly coital frequency, 1984			
M	9.1	6.6	.06[a]
SD	9.4	6.2	
n	51	49	
Monthly coital frequency, 1985			
M	8.7	7.7	.31[a]
SD	10.6	7.1	
n	44	45	
Contraceptive use score, 1984-88[b]			
M	2.1	2.4	.23[c]
SD	0.9	0.8	
n	32	34	
Contraceptive use score, 1990-94[b]			
M	2.2	2.3	.77[c]
SD	1.0	1.0	
n	37	24	
Contraceptive use score, 1984-96[b]			
M	5.0	5.4	.33[c]
SD	1.3	1.6	
n	24	18	

[a] By one-tailed *t* test; [b] higher scores indicate greater use;
[c] by two-tailed *t* test.

Of course, contraception is not always used properly. Compliance estimates may be attempted in clinical settings, using techniques such as return of unused portions of contraceptive products. In large-scale surveys such as the NLSY, such determinations are far less feasible. Based on the data available, it is not possible to say whether contraceptive failure may have contributed significantly to the differences in pregnancy rates between the two groups.

It serves to recall that all subjects in this analysis ranked in the upper quartile of general intelligence. In this context, it seems unlikely that improper use of contraceptives would account for a large share of the pregnancy differential between the two groups.

The high percentage of HMHM females who had never conceived prompted consideration of within-group differences. One study variable--body fatness--proved to be an interesting predictor. As a group, HMHM females who experienced a pregnancy by 1996 averaged more body fat in early adulthood their counterparts who never conceived. Mean 1981 BMI was 22.7 for the former group, compared to 20.6 for the latter (p = .11; two-tailed t test).

Libido. The dissertation model predicted that HMHM females would initiate intercourse at a later age than controls, but have higher libido thereafter. Libido was operationalized as self-reported sexual frequency for the month prior to interview in the 1984 and 1985 survey years.

Consistent with predictions, HMHM females reported higher mean coital frequencies than controls (Table 20). For 1984, the difference between groups was significant (p = .06). In 1985, mean frequencies were more similar between groups. HMHM subjects reported slightly more sexual activity, but differences fell substantially short of statistical significance (p = .31).

Further analysis suggested the 1985 results to be an artifact of the first NLSY downsizing. Coital histories of subjects whose participation was terminated in 1985 revealed a curious finding. Among the 11 HMHM subjects excluded as of 1985, mean monthly sexual frequency in 1984 had been notably higher than the group average (11.0 vs. 9.1). The pattern for terminated controls was precisely the opposite. Their mean coital frequency in 1984 had been substantially below their group's average (5.1 vs. 6.6).

For subjects who remained in the study, sexual frequency data for 1984 and 1985 were strongly correlated (r = .58, p < .001). Therefore, the 1985 results were likely biased by the elimination of the more active HMHM subjects and the less active controls. Because the 1984 results appeared more reliable, the prediction was considered supported.

The import of these findings extends beyond their support for a specific prediction. They must also be viewed in the context of other outcomes. The picture that emerges shows HMHM females reporting more sexual activity and somewhat less contraceptive use than controls, yet conceiving far less frequently. This is a phenomenon in need of explanation.

Number of children. As predicted, HMHM females not only experienced fewer pregnancies, but also fewer births (Table 21). The pattern mirrored that of pregnancy histories. Differences in mean number of live births trended toward, but did not reach statistical significance, for the 1984 survey year. By 1988, a significant trend toward fewer births in the HMHM group was evident. It persisted through 1996.

Controls elected to terminate pregnancies more often than did HMHM females (Table 22). Of subjects interviewed in 1996, six HMHM and 12 controls had elected to abort at least one pregnancy. Only one HMHM subject, compared to seven controls, reported doing so more than once. The number of pregnancies terminated voluntarily was seven in the HMHM group and 23 among controls. Absent this greater use of abortion services by controls, the parity differential between the groups would have been higher still.

Table 21
Longitudinal Birth Histories for HMHM and Control Groups

Characteristic	HMHM	Control	1-tailed *p* value[a]
# children born by 1984			
M	0.3	0.4	.13
SD	0.6	0.7	
n	65	58	
# children born by 1988			
M	0.5	1.0	.009
SD	0.8	1.2	
n	51	51	
# children born by 1996			
M	1.2	1.6	.10
SD	1.5	1.3	
n	39	40	
Childless in 1996			
Yes	49%	22%	.01[b]
No	51%	78%	
n	39	40	

[a] By *t* test unless otherwise indicated; [b] by one-sided Fisher's exact test.

The proportion of HMHM females still childless in 1996 was extraordinary. Almost half (49%) had never given birth, compared to 22% of controls. This difference was highly significant--and among the most notable of all study findings. As discussed above, neither the contraceptive or coital histories of HMHM females were consistent with this outcome.

The lower birth rate of HMHM females may seem predictable in light of their lower pregnancy rate. However, in the course of analysis, a second--and completely unexpected factor--emerged. HMHM females were far more prone to involuntary pregnancy loss than controls.

Results from the 1984 survey year were striking. Of HMHM females who had conceived at least once, almost half (47%) had experienced a miscarriage or a stillbirth. Among controls, the comparable rate was only 8% (Table 22). Of all pregnancies conceived by HMHM females as of 1984, fully 24% were lost. For controls, the corresponding figure was only 5%. These outcomes represented a significant difference between the groups (p = .02 by two-sided Fisher's exact test).

Table 22
Pregnancy Loss and Termination among HMHM and Control Groups[a]

Characteristic	HMHM	Control	p value
Had miscarriage by 1984			
Yes	47%	8%	.004[b]
No	53%	92%	
n	19	26	
Had miscarriage by 1988[c]			
Yes	35%	10%	.04[b]
No	65%	90%	
n	20	31	
Had miscarriage by 1996[c]			
Yes	29%	12%	0.16[b]
No	71%	88%	
n	21	33	
# elective abortions by 1996			
M	0.3	0.7	0.12[d]
SD	0.6	1.1	
n	21	33	

[a] Never-pregnant subjects excluded from analysis; miscarriage includes stillbirths; [b] by two-sided Fisher's exact test; [c] decreased percentage since 1984 reflects changes in composition of active sample-- see text; [d] by two-tailed t-test.

The explanation for this startling outcome was elusive. Certainly, it was not advanced age. When first asked about pregnancy loss, subjects were 19 to 27 years old. Thus, this was the maximum age at loss for miscarriages and stillbirths reported during the 1984 survey year.

Cigarette use may have been a factor among the affected controls; both had smoked daily since age 15. However, among HMHM females who had lost a pregnancy as of 1984, it was unclear that any used tobacco. Subjects had been asked At what age did you start smoking daily? Five of nine affected HMHM females described the question as inapplicable. Data for the other four were missing.

Drug abuse seemed an equally unlikely explanation. Both control subjects who had miscarried as of 1984 acknowledged use of marijuana or hashish 40-99 times in their lives. By comparison, four of the nine affected HMHM females reported never having used these drugs. The remaining five confirmed very light use (less than 10 occasions).

When asked about lifetime use of cocaine in 1988, six of the nine HMHM females again denied even a single use. One acknowledged using the drug once or twice; no data were available for the remaining two HMHM subjects. By contrast, one of the two affected controls admitted to using cocaine 100 or more times. The other denied any use.

The high rate of miscarriage or stillbirth among HMHM females persisted over the next 12 years of the NLSY. The percentage of affected HMHM females trended lower in 1988 and 1996, but was likely another artifact of NLSY downsizing. Several females who had miscarried as of 1984 were subsequently dropped from the study, contributing to a decline in the percentage of subjects reporting a pregnancy loss in later years. Regardless, differences between groups remained statistically significant through 1996.

Moreover, the percentage of total pregnancies that ended in miscarriage or stillbirth continued to diverge between the groups. Of pregnancies reported by HMHM females between 1985 and 1994, 23% ended in miscarriage or stillbirth--compared to 7% for controls (p = .10 by two-sided Fisher's exact test).

Care must be taken to understand these findings in accurate context. The data clearly indicate that an elevated incidence of miscarriages and stillbirths among HMHM females reduced the number of children born *to the group*. However, the analysis did not identify miscarriage or stillbirth as markers of infertility *in individuals*.

For example, nine HMHM females had miscarried or experienced a stillbirth as of 1984. Of these, seven remained active in the study in 1988. As of that survey year, each reported at least one successful birth. Four had borne two children, and one had given birth to three. Miscarriages and stillbirths did occur among other subjects in later study years. As of 1996, however, only one HMHM female who had experienced such an event remained childless.

In other words, childless subjects were rarely those who had miscarried or had a stillbirth. Rather, they were those who had not conceived at all. In 1996, 19 of 39 females remaining in the HMHM group were childless, as were 9 of 40 controls. Of the childless HMHM females, 17 had never been pregnant. This was also the case for six of the nine childless controls.

In 1979, subjects answered a question about their childbearing aspirations. At that time, only two of the childless females in each group expressed a desire not to have children. This raises the question of whether the high rate of childlessness among HMHM females was entirely voluntary.

To shed light on this question, data on contraceptive use were examined for all childless females. In the HMHM group, only two reported consistent use. A large percentage reported not using birth control (despite sexual activity) during multiple study years. More specifically, non-use of birth control was reported at least three times (and as many five) by 9 of the 19 females. Seven of the 19 answered no regarding contraceptive use during at least three consecutive interviews. One of the seven had conceived once, but terminated the pregnancy for reasons unknown.

Among childless controls, three of nine gave negative answers about contraceptive use for four of the seven queries incorporated into this analysis. Thus, it appears that a few controls, as well as a larger number of HMHM females, may have been attempting pregnancy without success.

Unfortunately, participants were never asked whether they had experienced difficulties in conceiving or had sought treatment for infertility. Accordingly, one can only speculate. Given that most of the childless females had expressed a desire for children--as well as evidence that many had not used contraceptives on multiple occasions--the possibility of fertility problems warrants serious consideration.

Within the HMHM group, childless females were about 18 months younger than those with children. No such age difference existed among childless and parous controls. However, all but 5 of the 19 childless HMHM females were using contraception at 1996 interview. In this context, a dramatic rise in pregnancy rates among these females seems unlikely regardless of their slightly younger age.

Certainly, fertility outcomes in both groups should be reassessed at the completion of their childbearing years. Nonetheless, it seems likely that HMHM females will maintain an unusually low birth rate. Such an outcome would be consistent with both the dissertation model and GTC theory.

Summary of Findings

The dissertation model made predictions about 14 variables. Each was classified as somatic, reproductive, or psychological in nature (Figure 1). Table 23 summarizes the status of the model after testing.

The most noteworthy outcome was the magnitude of support for the somatic and reproductive predictions. Though data for one (handedness) was considered inconclusive, the remaining eight were supported. Psychological predictions fared poorly by comparison.

Simply stated, these findings suggest that what sets HMHM females apart from controls are characteristics more readily described as biological than psychological.

The Relevance of General Intelligence

Individuals who scored in the top decile of arithmetic reasoning or mathematics knowledge were ineligible to serve as controls in this study. This standard insured that controls had none of the specific cognitive strengths that define the HMHM ability pattern. However, it also created differences in general intelligence between the two groups.

Table 23
Summary of Predictions and Outcomes

Relative to controls, HMHM females will	Supported	Not supported	Inconclusive
Be taller	√		
Be leaner	√		
Be more physically active	√		
Be more likely to be left-handed			√
Be older at menarche	√		
Be older at first intercourse	√		
Have fewer pregnancies	√		
Have fewer children	√		
Have higher libido	√		
Show less religiosity	√		
Have higher self-esteem		√	
Have less traditional role attitudes			√
Have fewer depressive symptoms		√	
Have a more internal locus of control		√	

This outcome was unavoidable because two of the three tests that define the HMHM pattern contribute significantly to the AFQT score. As already discussed, the AFQT is calculated from scores on four of the ten ASVAB subtests. It is a good indicator of *g*, the general factor of intelligence (Jensen, 1998). The four subtests have *g* loadings ranging from .84 for mathematics knowledge to .94 for word knowledge (Ree & Carretta, 1994). Instruments commonly marketed as IQ tests have *g* loadings in the range of .80 or above (Jensen, 1998).

AFQT scores are expressed as percentiles. Mean AFQT scores for HMHM and control females were 94.7 and 79.4, respectively--a difference of about 15 percentile points. These means were significantly different ($p < .001$; two-tailed t test) and necessitated further analysis.

The most obvious question was whether the HMHM ability pattern is simply a marker for general intelligence. Even if it were not, one had to ask whether the variables that set HMHM females apart from controls were linked--at least in part--to differences in general intelligence.

To answer the first question, a frequency analysis was performed to identify NLSY females with AFQT scores in the top decile. It yielded 287 females from a pool of 5,811. This made clear that HMHM and high IQ were far from synonymous; only 57 of these 287 females qualified for the HMHM group.

These results also revealed that only 1 in every 5 females who scored in the top decile of general ability had the HMHM pattern. Ten HMHM females had verbal scores that kept them below the top decile of the AFQT. Thus, the HMHM pattern represents a construct that differs from general intelligence.

Examining general ability as a confounding factor in the outcome of the dissertation model required a more elaborate analysis. Toward this end, a second comparison group consisting of females with high AFQT scores was created. The criteria for this group, abbreviated as HIIQ, were:

* AFQT score greater than or equal to the 90th percentile
* ineligibility for the HMHM group
* mechanical comprehension score below 52

A note about the latter criterion is necessary. Control subjects used to test the dissertation model had mechanical comprehension (MC) scores less than 48. Ideally, the HIIQ group would not have differed on this count. However, general ability influences outcome on all ASVAB subtests to varying degrees. Drawing a comparison group of sufficient size from a pool of highly talented females therefore required a minor increase in the maximum MC score. Though slightly higher than the uppermost score used to select the original control group, the maximum for the HIIQ group was nonetheless two full deciles below the minimum MC score required for the HMHM designation.

Table 24 compares the HMHM and HIIQ groups on measures of general and verbal ability. General ability was essentially identical between groups. HIIQ females trended toward better performance in verbal domains. However, this finding was eclipsed by more dramatic differences in technical domains, general science, and arithmetic reasoning.

Table 25 shows the magnitude of these differences. All favored the HMHM group and were consistent with their tendency toward cognitive strengths more typical of males than females.

Table 24
General and Verbal Scores of HMHM and HIIQ Females

ASVAB Measure	HMHM (n = 67)	HIIQ (n = 51)	2-tailed p value[a]
AFQT (%tile)			
M	94.7	94.4	.65
SD	4.3	2.8	
Paragraph comprehension			
M	58.8	59.6	.08
SD	2.4	2.5	
Verbal composite			
M	59.7	60.2	.17
SD	2.0	1.4	
Word knowledge			
M	59.6	59.9	.38
SD	2.2	1.3	

[a] By t test.

The pattern also shows that this high technical ability was associated with advantage in arithmetic reasoning. This is consistent with the view that spatial-mechanical aptitude enhances mathematical skill. The slightly poorer outcome of HMHM females in verbal domains also suggests a trade-off between verbal and nonverbal ability when general intelligence is held constant.

HMHM vs. HIIQ in the Dissertation Model

Comparing HMHM and HIIQ groups would have been an obvious way to examine whether life histories varied when specific cognitive strengths differed, but general intelligence did not. However, cognitive abilities were more similar among these two groups than among HMHM and control females. As a result, a sizable pool of subjects--or more sensitive testing instruments--would have been needed to detect statistically significant differences between the groups. The NLSY database lacked sufficient subjects for such an analysis.

Although inadequate for hypothesis testing, the data were sufficient for a descriptive study. Its purpose was to allow insight into whether differences in general intelligence confounded the results obtained with HMHM and control females. HMHM, HIIQ, and control females were included in this follow-up analysis. Group means were computed for the nine variables of the dissertation model affirmed in the primary analyses. Comparisons were also made regarding rates of miscarriage and childlessness. Results were translated into a series of bar graphs.

Table 25
Mathematical and Technical Scores of HMHM and HIIQ Females

ASVAB Subtest	HMHM (n = 67)	HIIQ (n = 51)	2-tailed p value[a]
Arithmetic reasoning			
M	64.9	63.9	.002
SD	1.2	1.8	
Auto and shop information			
M	54.5	47.1	<.001
SD	4.9	5.6	
Electronics information			
M	58.4	52.4	<.001
SD	5.9	6.0	
General science			
M	60.7	57.0	<.001
SD	4.6	5.8	
Mathematics knowledge			
M	65.9	65.3	.18
SD	1.6	2.5	
Mechanical comprehension			
M	64.9	47.0	<.001
SD	2.1	3.0	

[a] By t test.

A caution about these graphs is in order. Their purpose is to show relative rankings. Therefore, with the exception of Figure 13, the graphs do not have a zero-point scale. As a result, differences between groups may appear larger than is actually the case. Such impressions can be avoided by noting the scale, which is individualized to each graph.

Somatic traits. Figures 4 through 6 compare mean values for height, body mass index, and physical activity across the three groups. With the exception of height, HIIQ females were consistently intermediate to HMHM and controls. For height, such a gradient was absent; HIIQ were fractionally shorter than controls. Inasmuch as the difference between HIIQ females and controls amounted to only .05 inches, mean height among the latter two groups was essentially identical.

To allow for the largest sample, BMI results for 1986 were computed and graphed. Examination of BMI data for 1990 revealed no change in the fundamental pattern. In 1996, HIIQ females had a lower mean BMI than the HMHM group. Closer analysis again suggested an artifact of NLSY downsizing. The mean BMI of HMHM subjects dropped from the NLSY after 1990 had been modestly higher than that group's average (24.0 vs. 23.4). In the

HIIQ group, however, discontinued participants had averaged a much higher BMI than those retained (29.4 vs. 24.2). The pattern shown in Figure 5 is therefore likely to be the most accurate.

Reproductive histories. For ages at menarche and first intercourse, mean values for HIIQ subjects again fell between HMHM females and controls (Figures 7 and 8). By contrast, libido did not differ between HIIQ and controls. Both groups had mean values well below that of the HMHM group.

A gradient across the groups was also seen for number of pregnancies, with controls > HIIQ > HMHM (Figure 10). Nonetheless, number of births was similar for HIIQ and controls (Figure 11). Use of elective abortion accounted for this result. As of 1996, controls averaged three times as many voluntary abortions as HIIQ females. Regardless, both HIIQ and control females reported higher parity than the HMHM group.

The percentage of each group still childless as of 1996 also followed a gradient, with rates highest among HMHM females and lowest in controls (Figure 12). The rate of childlessness for the HIIQ group was intermediate, but not equidistant between HMHM and controls. Rather, it was closer to the lower rate of the latter group.

Owing to the dramatic findings regarding miscarriage and stillbirths among HMHM females, data for both 1984 and 1996 are presented (Figures 13 and 14). Never-pregnant females were excluded from these analyses.

At both endpoints, the percentage of HIIQ females who had experienced miscarriage or stillbirth was higher than among controls but lower than that of the HMHM group. The percentage of affected HIIQ females also trended higher with age. As a result, it was more similar to the experience of HMHM females in 1996 than in 1984.

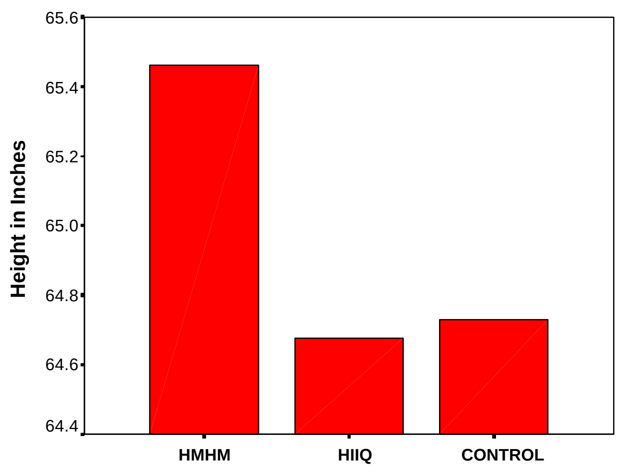


Figure 4. Adult height.

Mean adult height among HMHM (n = 54), HIIQ (n = 46), and control (n = 52) ability groups.

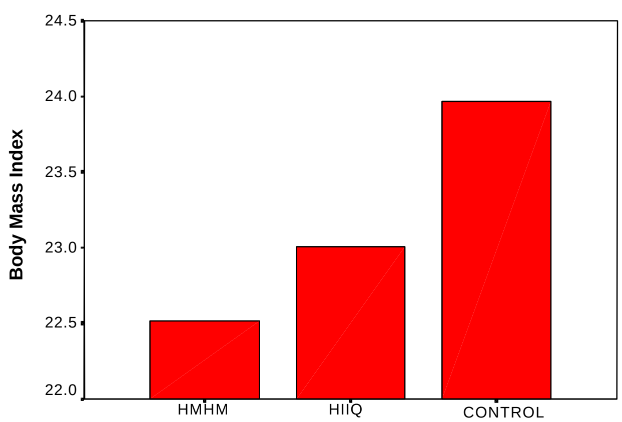
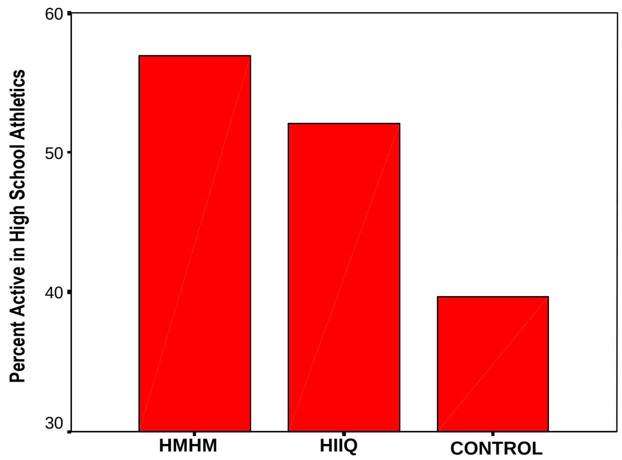


Figure 5. Body mass index, 1986.

Mean body mass index among HMHM (n = 52), HIIQ (n = 43), and control (n = 50) ability groups in 1986.



<u>Figure 6.</u> Physical activity.

Percentage of HMHM *(n* = 65), HIIQ *(n* = 48), and control *(n* = 58) ability group members who participated in high school athletic activities.

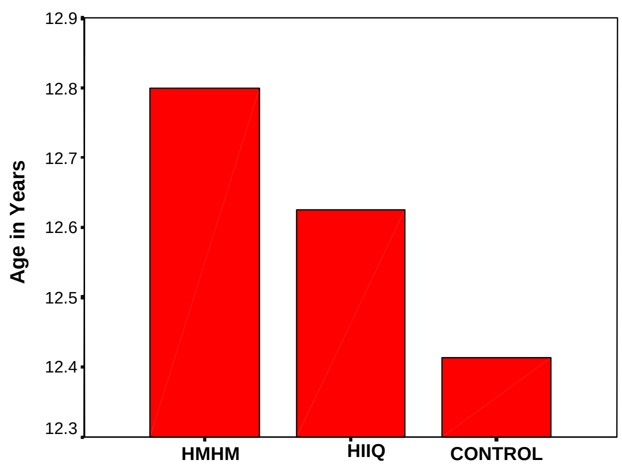
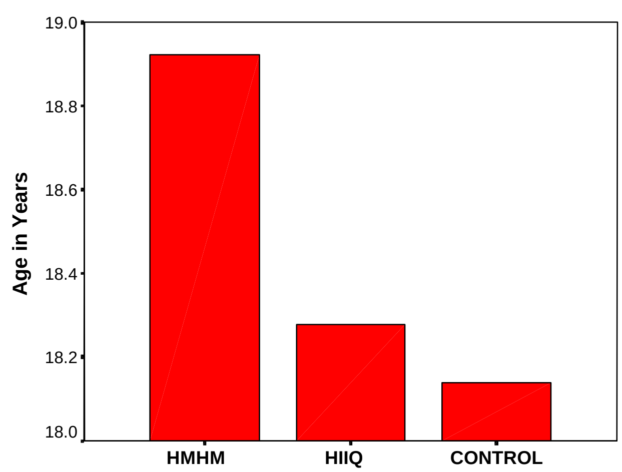


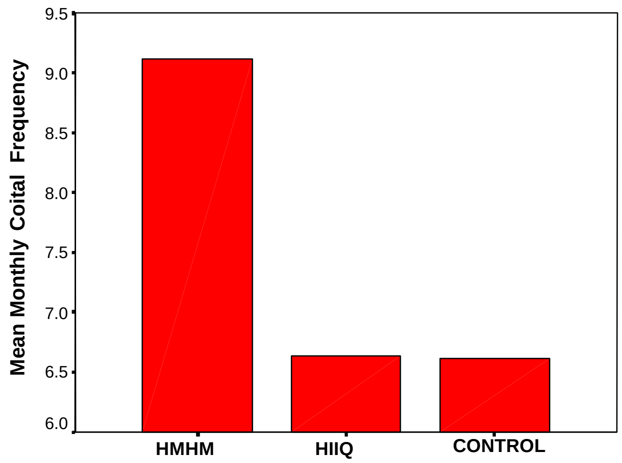
Figure 7. Age at menarche.

Mean menarcheal age for HMHM *(n* = 65), HIIQ *(n* = 48), and control *(n* = 58) ability groups.



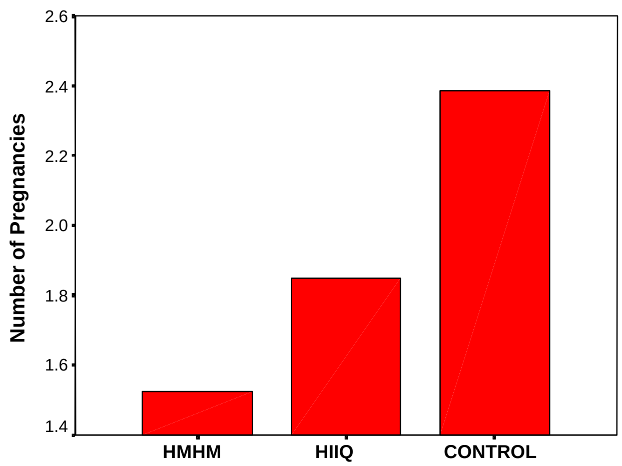
<u>Figure 8.</u> Age at first intercourse.

Mean age at first intercourse for HMHM *(n* = 52), HIIQ *(n* = 36), and control *(n* = 50) ability groups.



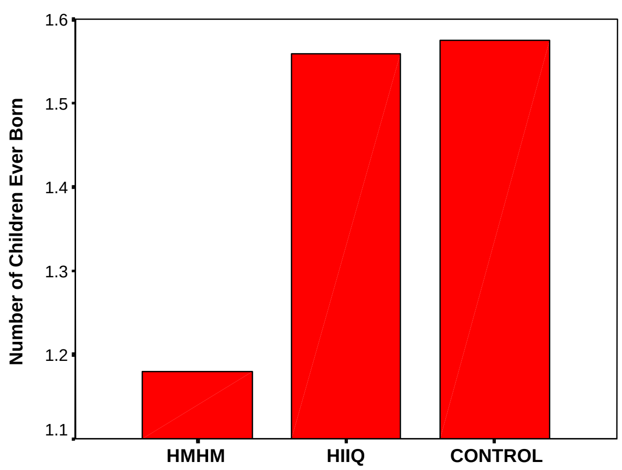
<u>Figure 9.</u> Sexual frequency, 1984.

Mean frequency of sexual intercourse in month prior to 1984 interview for HMHM (*n* = 51), HIIQ (*n* = 38), and control (*n* = 49) ability groups.



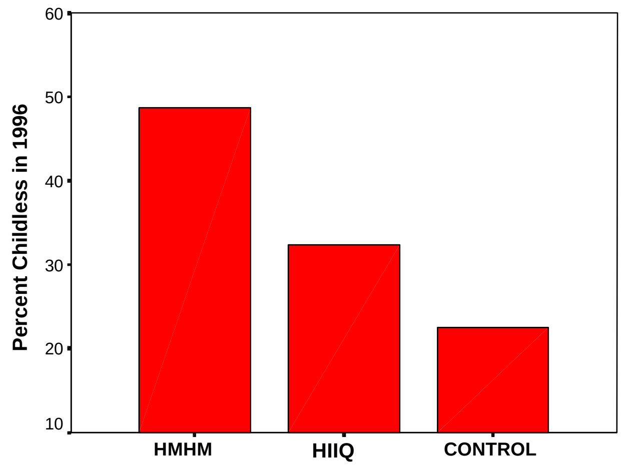
<u>Figure 10.</u> Pregnancies.

Mean number of pregnancies for HMHM (*n* = 38), HIIQ (*n* = 33), and control (*n* = 39) ability groups as of 1996.



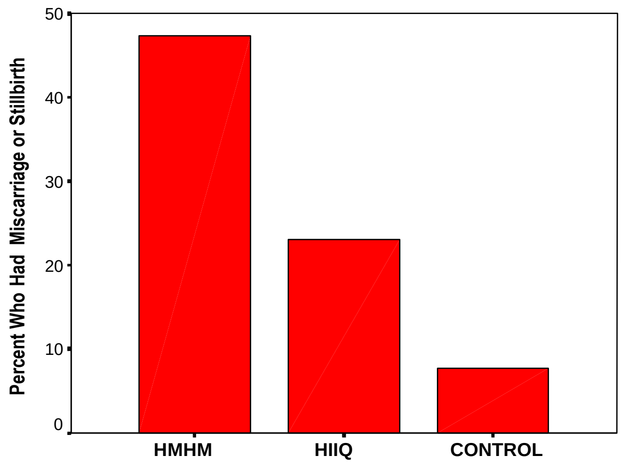
<u>Figure 11.</u> Parity.

Mean number of children ever born to HMHM (*n* = 39), HIIQ (*n* = 34), and control (*n* = 40) ability groups as of 1996.



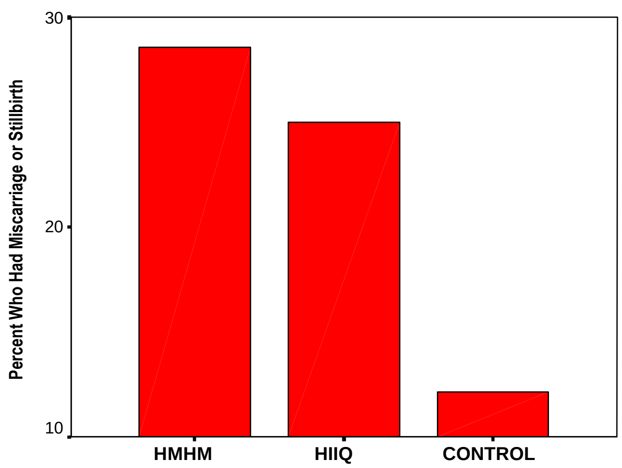
<u>Figure 12.</u> Childlessness.

Percentage of subjects who had never given birth to a child as of 1996 for HMHM *(n* = 39), HIIQ *(n* = 34), and control *(n* = 40) ability groups.



<u>Figure 13.</u> Pregnancy loss, 1984.

Percentage of subjects in HMHM *(n* = 19), HIIQ *(n* = 13), and control *(n* = 26) ability groups who experienced a pregnancy that ended in miscarriage or stillbirth prior to or during 1984.



<u>Figure 14.</u> Pregnancy loss, 1996.

Percentage of subjects in HMHM *(n* = 21), HIIQ (*n* = 24), and control (*n* = 33) ability groups who experienced a pregnancy that ended in miscarriage or stillbirth prior to or during 1996.

<u>Psychological</u> traits. Figure 15 shows mean scores on the NLSY religious attendance scale. Results revealed a fairly linear pattern of increasing attendance with the transition from HMHM to HIIQ to controls.

Summary. Life histories of HMHM females varied considerably from those of women matched for general intelligence (HIIQ) but lacking the specific characteristics of the HMHM pattern. In general, mean values on life history variables for the HIIQ group were intermediate to those of HMHM and control females (Figure 16). The consistency of this pattern--with HIIQ means intermediate to the other groups--was striking.

The most notable exception to this pattern was height. Although the HMHM group was tallest, means for HIIQ females and controls were almost identical. Findings relative to sexual frequency were similar; the HMHM group had the highest mean frequency, but HIIQ and controls did not differ. A gradient across the three groups was also absent in relation to parity, but was readily explained by use of elective abortion. Were it not for these voluntary terminations of pregnancy, the pattern of controls > HIIQ > HMHM would have remained.

The results suggest that all three patterns may have similar influences, with HIIQ and HMHM representing progressively higher (or lower) levels of exposure to common causal agent(s).

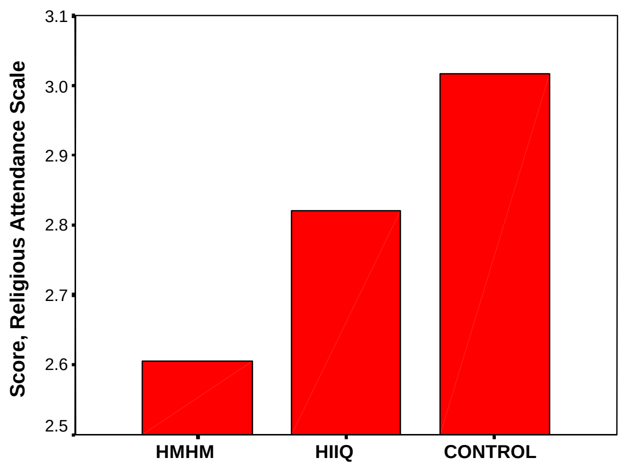


Figure 15. Religiosity.

Mean score in 1982 on NLSY religious attendance scale for HMHM *(n* = 66), HIIQ (*n* = 50) and control (*n* = 58) ability groups.

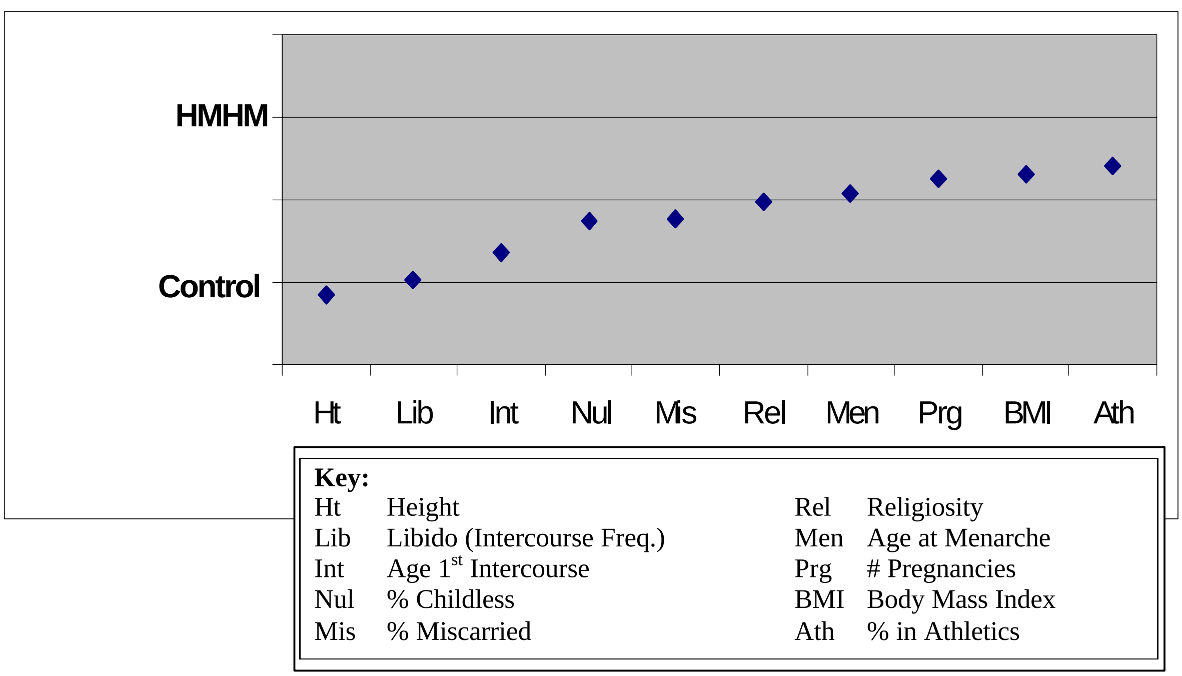


Figure 16. Summary of HIIQ group traits.

Mean values of HIIQ group shown relative to those of HMHM and control females for variables supported by the dissertation model--and for pregnancy loss and childlessness.

A Look at HMHM Males

The primary purpose of this dissertation was to study the correlates of the HMHM ability pattern in females. It was hoped, though not assumed, that doing so might yield clues as to why a pattern of mathematical and mechanical giftedness is rare among females.

Of all findings, the most provocative related to reproductive outcomes. As a group, HMHM females were remarkable for their unusually low birth rate and their high rates of childlessness and pregnancy loss. These results prompted interest in examining variables related to reproductive success among males with the same ability pattern.

For males, the number of relevant variables is naturally smaller. Moreover, NLSY interviewers restricted certain questions of potential interest to females. For example, interviewers did not ask males about their age at first intercourse or coital frequencies.

Because the HMHM pattern is much more common among males than females, assembling a sizable sample of HMHM males posed no problem. However, the standards used to select female controls proved unsuitable. Only eight NLSY males met the criteria. As an alternative, HMHM males were compared to men who scored in the top decile on the AFQT but did not fulfill criteria for both mathematical and mechanical giftedness. These comparison males are referred to as HIIQ-M.

To insure comparable levels of general ability, 15 males with the HMHM pattern who scored below the 81th percentile on the AFQT were excluded from the analysis. The resulting group of HMHM males is abbreviated as HMHM-M.

In 1980, the HMHM-M group had 310 members; the HIIQ-M had 210. These proportions differ dramatically from those found among NLSY females, where HIIQ outnumbered HMHM subjects by five to one. The average age of HMHM-M and HIIQ-M subjects was highly comparable. The mean month and year of birth was in the summer of 1959 for both groups.

Highlights of the marital and parenting histories of these males, as well as their family aspirations, revealed notable similarities. At their entry into the NLSY, large majorities of HMHM-M and HIIQ-M subjects expressed a desire to have families. Fully 93% of subjects in each group wanted children. The mean number of offspring desired was similarly identical, at 2.4 for both groups.

Interest in marriage was strong. Only 2% of HMHM-M and 4% of HIIQ-M subjects indicated a desire to remain single. This difference between groups was not statistically significant.

On follow-up in 1994 and 1996, some differences had emerged between the groups. Fifteen years after their initial interview, life histories showed that:

* As of 1994, 19% of HMHM-M and 26% of HIIQ-M subjects had never married. This outcome approached statistical significance, $\chi 2$ (1, n = 337) = 2.5, p = .11. By 1996, the gap had closed somewhat, with 16% of HMHM-M and 20% of HIIQ-M remaining in the never-married category.

* More than half (51%) of HIIQ-M had not fathered a child by 1994, compared to 40% of HMHM-M. This difference was significant, χ^2 (1, n = 339) = 4.2, p = .04. Two years later, this trend had moderated, with 41% of HIIQ-M and 33% of HMHM-M subjects still childless, χ^2 (1, n = 329) = 2.5, p = .11.

* As of the most recent survey year, HMHM-M subjects had fathered an average of 1.4 children. The corresponding mean for the HIIQ-M group was 1.1. This result was statistically significant (p = .04, two-tailed t test).

Tables 26 and 27 show the cognitive characteristics associated with these outcomes. Mean AFQT scores actually differed only by .01 percentile; the difference of .1 shown in the table reflects rounding.

Despite virtually identical AFQT scores, subtest scores revealed highly significant differences on all tests measuring technical, verbal, or quantitative aptitudes. HIIQ-M outscored HMHM-M on word knowledge, paragraph comprehension, and the ASVAB verbal composite. On all tests of mathematical and technical proficiency, however, HMHM-M outperformed HIIQ-M. This is consistent with the pattern of their female counterparts (Tables 24 and 25).

The data for these males--like that for HMHM and HIIQ females--suggest that at a given level of general ability, some degree of trade-off occurs between verbal and mathematical/technical strengths. The direction of this trade-off differs between sexes. Among highly intelligent females, verbal ability is overwhelmingly favored. Among their male counterparts, a pattern characterized by strength in mathematics and technical areas prevails by a ratio of 3:2. The question that begs to be answered is why.

Table 26
General and Verbal Scores of HMHM-M and HIIQ-M Groups

ASVAB Measure	HMHM-M (*n* = 310)	HIIQ-M (*n* = 214)	2-tailed *p* value[a]
AFQT (percentile)			
M	94.1	94.0	.87
SD	4.5	3.1	
Paragraph comprehension			
M	58.3	59.1	.003
SD	3.3	2.6	
Verbal composite			
M	59.2	60.1	<.001
SD	2.1	1.3	
Word knowledge			
M	59.1	60.0	<.001
SD	2.1	1.3	

[a] By *t* test.

Table 27
Mathematical and Technical Scores of HMHM-M and HIIQ-M Groups

ASVAB Subtest	HMHM-M (*n* = 310)	HIIQ-M (*n* = 214)	2-tailed *p* value[a]
Arithmetic reasoning			
M	64.9	64.0	<.001
SD	1.3	2.2	
Auto and shop information			
M	62.2	58.2	<.001
SD	5.7	6.8	
Electronics information			
M	63.7	61.4	<.001
SD	4.4	5.2	
General science			
M	63.7	62.6	.003
SD	4.2	4.1	
Mathematics knowledge			
M	66.1	65.0	<.001
SD	1.7	2.6	
Mechanical comprehension			
M	66.7	58.4	<.001
SD	2.2	5.2	

[a] By *t* test.

Chapter 5: Discussion

Making sense of a large number of variables requires a logical framework. A primary consideration is the relative importance of each; rarely will all be equally valuable for answering the questions at hand.

Certainly, this was the case here. Some variables were more strongly linked than others to the conceptual basis for the research model. In addition, the nature of the data available for testing the research predictions varied. Some data were in the form of well-established, easily measured concepts. Others reflected abstract constructs. Finally, while most variables were vulnerable to confounding factors, the likelihood or magnitude of such influences was not constant among them.

Recognition of these facts prompted a systematic approach that allowed variables to be weighed relative to their importance. Figure 17 depicts the approach in diagrammatic fashion.

On the horizontal axis, the term relevance of variable encompasses two considerations. The first is the strength of the relationship between the variable and Nyborg's theory. The second is the potential, relative to other variables, for outcomes to have been confounded by other influences. The vertical axis classifies variables based on ease of measurement and the quality of the data acquisition tool.

Applying this approach to the dissertation model makes clear that that height is a more primary variable than locus of control. Height is a concrete, easily measured concept; locus of control is abstract and more difficult to measure. Similarly, age at menarche yields more valuable information than libido because the former represents a single event. The latter is subject to change over the course of many years, yet was assessed only twice during the course of the NLSY.

Figure 18 shows the variables of the dissertation model classified by the described method. Notably, all six variables listed in the bottom row derive directly from GTC theory. All are concrete concepts, and several do not change with time. What separates a variable such as parity from those listed to its left is the greater potential for confounding. Age at menarche, for instance, is less susceptible to external influences than is age at first intercourse. In fact, the three variables classified as primary are recognized as substantially heritable (Allison et al., 1996; Loesch, Huggins, Rogucka, Hoang, & Hopper, 1995; Stunkard, Foch, & Hrubec, 1986).

This is not to say that variables listed in the left column are free from confounding. Distinctions represent a judgment as to the relative potential for bias. Similarly, variables listed in the upper right portion of the matrix are considered least critical to interpreting the results. These concepts are more weakly linked to GTC theory and/or more difficult to assess.

The discussion that follows focuses on how well the findings fit the dissertation model. As the underlying theory assumes a critical role of hormones, outcomes are discussed only in the context of endocrine influences. That non-hormonal factors also influence the variables is acknowledged. However, their role is outside the scope of this discussion.

The Primary Variables

Age at menarche, height, and body mass index are primary variables. Although the latter is subject to variation over time, the NLSY database provided sufficient information to allow for observation of such changes.

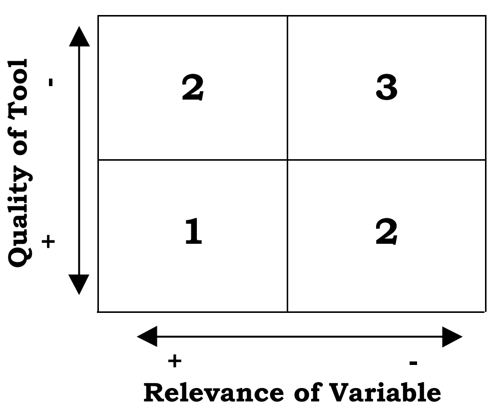


Figure 17. Schema for classifying variables.

Criteria for ranking importance of variables as primary (1), secondary (2), or tertiary (3).

2 handedness libido physical activity sex-role traditionality	**3** depressive symptoms locus of control religiosity self-esteem
1 age at menarche body mass index height	**2** age at first intercourse number of children number of pregnancies

Figure 18. Rankings of study variables.

Classification of study variables as primary (1), secondary (2), or tertiary (3) in accordance with the conceptual model shown in Figure 17.

Significantly, all predictions relative to these variables were affirmed. HMHM females were significantly older at menarche than controls and were also significantly taller. Moreover, while BMI increased with age in both groups, HMHM females were leaner than controls at all endpoints examined.

Is there a parsimonious explanation for the clustering of these traits--greater height, less body fat, and later menarche? As estrogens influence all three, the answer is yes.

Age at menarche. Even in their pre-pubertal years, girls who have an early menarche show higher circulating levels of estradiol than those who mature more slowly (Vihko & Apter, 1984). These higher levels persist through the third decade of life (Apter, Reinila, & Vihko, 1989)--if not longer.

Remarkably, an inverse association between recalled age at menarche and circulating levels of estradiol, bioavailable estradiol, and estrone was recently reported in a cohort of postmenopausal women (Madigan et al., 1998). In fact, later menarche is considered [one of the] markers of lower lifetime exposure to circulating ovarian hormones (Rockhill, Moorman, & Newman, 1998). The notion that HMHM females had lower lifetime estrogen exposure than controls is consistent with this view.

Body mass index. The consistently lower BMI of the HMHM group also yields to this interpretation. A positive association with body mass index (or other measure of adiposity) repeatedly has been reported for plasma levels of both estradiol and estrone (Bruning et al., 1992; Cauley, Gutai, Kuller, LeDonne, & Powell, 1989; Kaye, Folsom, Soler, Prineas, & Potter, 1991; Madigan et al., 1998).

In the Nurses Health Study, the correlation between estrogenic hormones and BMI ranged from an r value of .37 for estrone to .63 for estradiol (Hankinson et al., 1995). These values reflect adjustments for co-variates linked to risk of breast cancer: age, height, smoking, and alcohol use.

Reliance on these data can be criticized on grounds that most studies involved postmenopausal women. After menopause, plasma estrogen levels no longer reflect ovarian output, but rather the continuing influence of factors such as adipocytes, exercise, and alcohol intake. However, an association between BMI and estrogen fractions has been shown in pre-menopausal women (Bruning et al., 1992).

Admittedly, a larger body of evidence would be desirable, as BMI data in this study were for women of childbearing age. Still, the ultimate source of hormones is probably not critical. It is doubtful that estradiol of ovarian origin differs in effect from the same substance derived from other mechanisms.

In fact, a recent investigation documented an inverse relationship between body fat and age at menarche among pre-menopausal females (Kirchengast, Gruber, Sator, & Huber, 1998). The results are compelling because dual-energy x-ray absorptiometry-a highly precise form of measurement--was used to quantify body composition.

In short, the findings regarding BMI in HMHM females are consistent with lesser exposure to estrogens. Moreover, these data are similar to previous reports regarding the relationship of age at menarche to body fatness.

Height. The greater height of HMHM females also suggests lesser exposure to estrogens. GTC theory predicts that high levels of estrogens will result in shorter stature because intense hormonal stimulation encourages early closing of the epiphyses of the long bones (Nyborg, 1994). Once closed, further skeletal growth cannot occur. This concept is hardly speculative; tall stature in females is treated with high doses of estrogen to encourage closure of the epiphyses (Binder, Grauer, Wehner, Wehner, & Ranke, 1997).

Several investigators have found inverse associations between plasma estrogens and adult height (Madigan et al., 1998; Yoo et al., 1998). Also noteworthy are positive correlations between menarche and adult height (Georgiadis et al., 1997; Helm, Munster, & Schmidt, 1995; Shangold, Kelly, Berkeley, Freedman, & Groshen, 1989). Such findings are consistent with GTC theory as well as with the results here.

Nonetheless, the literature also contains contrary data. A positive correlation between height and plasma estradiol during the follicular phase of the menstrual cycle has been reported (Dorgan et al., 1995). Moreover, reported correlations between height and risk of breast cancer (Albanes & Taylor, 1990; Swanson et al., 1989; Vatten & Kvinnsland, 1990) raise the possibility that estrogens are positively, rather than negatively, associated with height.

Such a relationship remains speculative. No physiological basis for a relationship between height and estrogens is known (Dorgan et al., 1995). Notably, the Nurses Health Study found no relationship between height and plasma estrogens (Hankinson et al., 1995). Moreover, other studies have found no relationship between breast cancer and height at ages 10-14 (Le Marchand, Kolonel, Earle, & Mi, 1988) or a negative association between height and risk of the disease (Parazzini et al., 1990).

The positive correlation between height and estradiol reported by Dorgan may well be an anomaly. Another possibility is that the relationship is nonlinear (Nyborg, 1994). Within the normal range of values, both low and above-average levels may encourage growth, while abnormally low and very high levels inhibit it. Alternatively, the influence of estradiol on height may depend on its concentration relative to other hormones.

Based on current information, it is reasonable to consider lesser exposure to estrogens a likely factor in the greater height of HMHM females.

Secondary Variables Related to Reproduction

Seven characteristics listed in the dissertation model are classified as secondary variables. Of these, four are reproductive variables and will be addressed here. The remaining three--handedness, physical activity, and traditionality--will be discussed in the next section.

Age at first intercourse. HMHM females were significantly older than controls at first sexual encounter. This observation is consistent with their later menarcheal age. A number of studies report positive associations between early sexual debut and early menarche, or viceversa (Andersson-Ellstrom et al., 1996; Helm & Lidegaard, 1990; Larsen & Kragstrup, 1997). The obvious implication is that initiation of coitus is influenced, at least in part, by the rate of pubertal development.

Common markers of maturational rate include stages of breast development, menarche, and figure changes. All are recognized as estrogen-dependent (Apter & Vihko, 1985; de Ridder et al., 1990; de Ridder et al., 1992). However, age at first intercourse is not so easily categorized.

Social variables obviously have considerable potential to confound coital debut; the same cannot be said for a variable such as age at menarche. Further, the relationship between age at first intercourse and estrogen status is less than clear. Estradiol increases female sexual initiation in nonhuman primates (Zehr, Maestripieri, & Wallen, 1998), but a similarly direct relationship in humans has not been reported.

Nonetheless, estrogen exposure can be viewed as an indirect influence on age at first intercourse--one that acts via effects on overall pubertal development. In this sense, the later age at sexual debut of HMHM females fits with other variables suggesting lesser estrogen exposure during adolescence and adult life.

Androgens also influence pubertal development (Longcope, 1986; Vanderzanden, 1989). Axillary and pubic hair, as well as acne, are familiar manifestations (Rebar, 1988). The possible role of androgens in the reproductive outcomes of HMHM and control females is considered below.

Number of pregnancies. As noted earlier, confounding factors can influence variables such as pregnancy and parity. Contraceptive use seemed a logical explanation for the significantly lower number of pregnancies among HMHM females. However, these subjects were as likely as controls to report periods of sexual activity without contraceptive use. In addition, their stated desire for children was virtually the same.

A recent longitudinal study of reproduction in Austrian women offers an insight worthy of consideration. In the context of his findings, the investigator concluded:

> [M]arked associations between female body build, above all body fat development, and hormonal induced events such as menarche, menstrual cycle patterns, fertility, and menopause were observed. *A more feminine type of body build combined with moderate development of body fat was associated with earlier menarche, more regular menstrual cycle patterns, higher reproductive success, and a delayed menopause* (Kirchengast, 1998, p. 251); (emphasis added).

These observations fit the facts here as well. Control females experienced menarche earlier, had a higher body mass index, and better fertility.

The relationship between earlier menarche, moderate body fatness, and fertility is well established (Frisch, 1987). A number of mechanisms account for the phenomenon (Frisch, 1990). Adipose tissue contributes to the estrogen pool, independent of ovarian output. Body fatness also influences the nature of estrogen metabolism, with the result that leaner women tend to produce less potent variants of the hormone. In addition, fat tissue can convert androgens to estrogens.

As a consequence, females of moderate body fatness have more regular ovulatory cycles--and hence, enhanced fertility (Frisch, 1990; Nyborg, 1994). Investigators who study nonhuman primates have found the same relationship. Their research has found, for instance, that fatter females have a shorter interval to pregnancy than their leaner counterparts (McFarland, 1997).

The field of assisted reproduction has provided further insight into the relationship between estrogens and fertility. Recently, investigators reported that blood estradiol levels on the fourth day of the menstrual cycle predicted the success of drug-induced ovarian stimulation (Phelps et al., 1998). Patients with higher estradiol levels (> 75 pg/mL) produced more oocytes in response to ovulation induction and were much more likely to actually achieve pregnancy. The clinical pregnancy rate was 42% for the high-estradiol group, compared to 9% among those with lower estradiol levels.

Among women undergoing artificial insemination, chances of conception were lower among women who were very lean or very obese (Zaadstra et al., 1993). These findings suggest a curvilinear relationship between estradiol and fertility. Alternatively, the ratio of estradiol to other hormones may be critical.

Estradiol may affect not only ovulation, but also the characteristics of the oocyte. For fertilization to occur, the sperm must penetrate the zona pellucida (ZP), the membrane surrounding the egg. In an analysis of embryos from women undergoing in vitro fertilization (IVF), investigators found that patients with unexplained infertility had significantly thicker zonae than those whose difficulties were attributed to endometriosis or male-factor infertility (Loret De Mola, Garside, Bucci, Tureck, & Heyner, 1997).

Naturally, this raises the question of whether the thickness of the zona actually influences the outcome of IVF. A highly significant relationship between zona thickness and fertilization has been observed; ova that remained unfertilized after exposure to normal sperm were thicker than those that were successfully penetrated (Bertrand, Van den Bergh, & Englert, 1995).

Further work has examined the relationship between a patient's maximum estradiol level and the characteristics of her oocytes. A significant, linear, decreasing relationship was found between a patient's maximum hormone level and the mean ZP thickness of her eggs (Bertrand, Van den Bergh, & Englert, 1996). Other laboratories report similar results. Specifically, the most fertilizable eggs are those from estrogen-rich pre-ovulatory environments (Botero-Ruiz et al., 1984; Kreiner, Liu, Itskovitz, Veeck, & Rosenwaks, 1987; Wramsby et al., 1981).

Clearly, estradiol influences fertilization rates in vitro. Arguably, this is prima facie evidence that its effects in vivo are similar. Actually, the point may well be moot. Even in the absence of such a role, evidence that estrogens modulate human fertility is abundant.

While ample evidence suggests that HMHM females differ from controls in their pattern of estrogen exposure, the possible role of androgens requires consideration as well. In females, both the ovaries and adrenal glands produce a variety of androgens. Fertility research has focused primarily on testosterone. Lesser-known androgens, such androstenedione, have been considered in some investigations.

Some intriguing results have been reported from a longitudinal study of Finnish females. Subjects ranged in age from 7 to 17 at entry. Hormone levels and reproductive outcomes were followed for 12 years. Testosterone and androstenedione levels at the inception of the study were highly correlated with those recorded more than a decade later (Apter & Vihko, 1990). Females who never conceived had significantly higher age-adjusted levels of both hormones than females who had experienced pregnancy. Nonetheless, neither hormone was related to sexual behavior or contraceptive use.

Physicians have long recognized a relationship between high levels of testosterone and reproductive abnormalities (Rebar, 1988). Menstrual irregularities are a common complaint among females who show clinical signs of hyperandrogenism, such as acne or hirsutism (Held, Nader, Rodriguez-Rigau, Smith, & Steinberger, 1984; Steinberger, Rodriguez-Rigau, Smith, & Held, 1981). Compelling evidence indicates that the relationship between testosterone and fertility problems reflects some degree of causality.

High levels of testosterone prolong the follicular phase of the menstrual cycle and shorten the luteal phase (Smith, Rodriguez-Rigau, Tcholakian, & Steinberger, 1979). These alterations adversely affect ovulation and fertility (Rantala, Stenman, & Koskimies, 1988). In addition, testosterone may inhibit fertilization via effects on the pH of the endocervical mucus (Jenkins, Brook, Sargeant, & Cooke, 1995). Fortunately, normalization of testosterone levels results in a high rate of successful pregnancies (Steinberger, Smith, Tcholakian, & Rodriguez-Rigau, 1979).

That HMHM females had fewer pregnancies than controls is insufficient to infer that androgen levels differed between the groups. In a correlational study such as this one, such a possibility would merit consideration only if supported by a wider range of outcomes than discussed thus far. Further evidence is also needed because studies linking androgens to infertility usually involved subjects who had never conceived. Only 45% of HMHM females had never been pregnant. In at least a few cases, this was obviously intentional.

Number of children. Fewer pregnancies among HMHM females naturally resulted in fewer births. However, the high rate of pregnancy loss among HMHM females exacerbated this outcome. This finding, which was completely unexpected, requires careful consideration within the framework of GTC theory.

The NLSY data set makes no distinction between miscarriages and stillbirths. This is understandable, since both represent loss of a pregnancy. Moreover, some discrepancies exist in the medical literature about the meaning of the terms.

Miscarriage is the more common form of pregnancy loss. In the medical literature, the term spontaneous abortion is generally preferred. An estimated 10% of all pregnancies end in miscarriage (Knudsen, Hansen, Juul, & Secher, 1991; Sandler, 1977). About 85% of these occur during the first trimester (Berkow, 1992).

Stillbirth--usually defined as the delivery of a dead fetus--occurs later in gestation. In the March of Dimes Preterm Prevention Trial, the term was reserved for pregnancy losses that occurred at 20 weeks of gestation or later. Analysis of outcomes for the 34,350 pregnancies followed in the trial put the frequency of stillbirth at 1% (Copper, Goldenberg, DuBard, & Davis, 1994)

Chromosomal abnormalities have been documented in 50-85% of fetuses lost to miscarriage (Boue, Bou, & Lazar, 1975; Cowchock, Gibas, & Jackson, 1993; Hassold et al., 1980). A host of maternal factors influence pregnancy loss as well. Advanced maternal age, smoking, substance abuse, Black race, and limited education have been acknowledged as risk factors (Copper et al., 1994; Feldman, 1992; Ogunyemi, Jackson, Buyske, & Risk, 1998; Parazzini et al., 1997). Yet, there was no indication that HMHM females who reported a pregnancy loss during or prior to 1984 had any of these characteristics. The group's high rate of pregnancy loss was therefore all the more remarkable.

Research in nonhuman primates has provided some intriguing evidence about pregnancy loss. Prospective study of fertility and pregnancy outcome in a group of pigtail macaques found that animals with less body fat prior to conception were more likely to experience spontaneous abortions or stillbirth (McFarland, 1997). While much research in humans confirms the importance of weight gain during pregnancy, a similar relationship between pre-pregnancy fat stores and avoidance of miscarriage or stillbirth remains to be corroborated.

By contrast, a large literature documents the importance of minimum levels of body fat for regular ovulation and conception (Frisch, 1987; Frisch, 1990; McFarland, 1997). In addition, weight gain during pregnancy and infant birth weight are influenced by a woman's pre-pregnancy weight (Abrams, Carmichael, & Selvin, 1995; Bakewell, Stockbauer, & Schramm, 1997; Eastman & Jackson, 1968; Ricalde, Velasquez-Melendez, Tanaka, & de Siqueira, 1998). Both are important determinants of infant survival.

The notion that fat stores prior to conception influence the likelihood of delivering a healthy infant makes intuitive sense. However, such a finding would apply here only if HMHM females who lost a pregnancy were in a range associated with increased risk. Regrettably, the NLSY data are not detailed enough to permit estimation of body fat immediately prior to ill-fated pregnancies. Accordingly, the lower BMI of HMHM females remains a plausible, yet speculative, basis for explaining their unusually high rate of pregnancy loss.

Body fatness and hormone levels are often related--although not necessarily so. Relationships between hormone status in the early phases of pregnancy and ultimate outcome have long been recognized. Among women with high-risk pregnancies, an abnormally low level of estradiol strongly predicts pregnancy failure (Kunz & Keller, 1976). Similarly, women who lose pregnancies have lower blood levels of estradiol during the early weeks of gestation than those who deliver normally (Aksoy, Celikkanat, Senoz, & Gokmen, 1996). Earlier findings regarding height, BMI, and age at menarche strongly suggest that HMHM females have lower lifetime exposure to estrogens. Whether these presumably lower levels would have been maintained during their pregnancies cannot be determined.

The ratio of free testosterone to total testosterone may also predict pregnancy outcome. In one study, all women with ratios of 1.05 or higher miscarried (Aksoy et al., 1996). By contrast, all patients with ratios below .85 experienced normal pregnancy.

Investigators have observed a similar trend in a Japanese population. Although the cut-off values varied somewhat, the pattern was the same. Women who had a high ratio of free testosterone to total testosterone during their first trimester consistently aborted; those with a low level did not (Takeuchi, Nishii, Okamura, Yaginuma, & Kawana, 1993). Of women with ratios in the intermediate range, about one in four aborted.

Given the multi-factorial nature of miscarriage and the lack of any specific medical information about the pregnancy losses among HMHM females, contemplating a role of free testosterone in these outcomes may seem wildly speculative. What makes the suspicion somewhat less so is the sexual history of these females.

In 1984, HMHM females reported significantly more sexual activity than controls. HMHM females who had miscarried were unusually active, reporting an average sexual frequency of 14 times per month. This exceeded the mean frequency of remaining group members by 66%. As discussed below, it is consistent with recognized effects of free testosterone on sexual motivation.

Though it may seem that free testosterone might be as important as estrogen in explaining differences between HMHM females and controls, this inference is best resisted. It is not possible to determine when and if free testosterone is an independent variable. During pregnancy, for example, the ratio of free testosterone to total testosterone is strongly influenced by sex-hormone binding globulin (SHBG). SHBG binds to almost all of the testosterone in plasma, rending it inactive (Takeuchi et al., 1993). Free testosterone is the tiny fraction left untouched by SHBG; only it retains biological activity.

However, multiple interactions are at work here. Sex hormones---particularly estradiol--often cause dramatic increases in SHBG (Casson et al., 1997; Mathur, Landgrebe, Moody, Semmens, & Williamson, 1985). Thus, a high level of free testosterone may simply be a marker for inadequate levels of SHBG. Such a deficit, in turn, may reflect little more than estradiol levels too low for optimal reproductive outcome.

Complicating the picture is that estradiol and SHBG do not always correlate in this way. In the Japanese study cited above, SHBG correlated positively with estradiol--and negatively with free testosterone--in women who delivered normally. In the women who aborted, however, bioavailable testosterone was not significantly related to SHBG or estradiol.

In nonpregnant women, ovarian hormones reportedly show no influence on adrenal output of androgens (Azziz et al., 1998). This finding is notable because, in females of childbearing age, both the adrenal glands and the ovaries contribute to the total pool of androgens. However, the number of women studied was small, and most subjects were overweight.

Thus, the question that remains is whether free testosterone is independent of estrogen status in some females but not others. To date, high levels of free testosterone coupled with normal levels of estradiol remain a phenomenon observed primarily in obese women. The underlying mechanism appears to be an adverse effect of adipose tissue on levels of SHBG (Zumoff, 1988). The effect appears stronger in women with upper-body obesity than in those with excess fat concentrated in the hips and lower body (Bernasconi et al., 1996). However, few HMHM females were obese.

With so few specifics about the pregnancy losses of HMHM females, the precipitating factors cannot be determined. However, hormonal influences do appear to be a plausible explanation--particularly in light of their young age, absence of other risk factors, and overall reproductive histories.

This possibility was strengthened recently when investigators examined the incidence of chromosomal abnormalities in fetuses that had been spontaneously aborted. Chromosomal abnormalities were far less common in fetuses that had been carried by women with a history of menstrual irregularities (Hasegawa, Takakuwa, & Tanaka, 1996). Thus, for a subset of women, hormonal influences may be a common factor in both pregnancy loss and related reproductive disorders.

Libido. The dissertation model predicted that HMHM females would have a higher libido than controls despite a later age at sexual debut. At first glance, these concepts may seem contradictory. An understanding of their theoretical basis shows otherwise.

The prediction regarding age at first intercourse derived from the expectation, now confirmed, of a slower maturational tempo in HMHM females. The notion that their libido would ultimately be stronger than that of controls stemmed from a conceptualization of these women as having a more androgynous or masculine profile; greater libido would be consistent with such (Jones & Barlow, 1990; Nyborg, 1988; Wilson & Lang, 1981). This tendency, however, would not be expected prior to full sexual maturation.

Using data for the 1984 survey year, the prediction of higher libido among HMHM females was supported. It again suggests differences in levels of bioavailable testosterone between HMHM and control females.

Androgens are widely recognized as primary influences on sexual arousability and excitement in both sexes (van Lunsen & Laan, 1997). Several studies show that marital coital frequency co-varies with the wife's level of free testosterone at mid-cycle (Morris, Udry, Khan-Dawood, & Dawood, 1987; Persky, Lief, Strauss, Miller, & O'Brien, 1978; van Lunsen & Laan, 1997). The relationship between bioavailable testosterone and sex behavior is further demonstrated by the longstanding use of exogenous testosterone to restore sexual function in females with hypoactive desire (Kaplan & Owett, 1993; Sherwin & Gelfand, 1987; Warnock, Bundren, & Morris, 1997).

Again, the primary question is whether free testosterone is an independent variable. As discussed above, it is simply not possible to say whether high levels of free testosterone are independent phenomena, or a natural consequence of lower levels of one or more estrogens.

Secondary Variables Independent of Reproduction

Physical activity. A higher percentage of HMHM females than controls participated in high school athletics. While this does not prove that members of the HMHM group were more active throughout life, it seems more likely than not. Levels of physical activity show some concordance throughout different stages of the lifespan (Malina, 1996; Telama, Leskinen, & Yang, 1996).

A more physically active lifestyle is consistent with the influences discussed so far. Physical activity in adolescent females has been correlated with lower blood levels of estradiol at mid-cycle (Nagata, Kaneda, Kabuto, & Shimizu, 1997). The delayed menarche and irregular menstruation of athletes similarly supports this inverse relation between estradiol and activity (Frisch et al., 1981), as do findings that physical activity predicts lower circulating levels of estrone (Cauley et al., 1989).

The greater proclivity of HMHM females to participate in sports is also consistent with higher levels of bioavailable testosterone. Females with higher levels of free testosterone are more inclined toward intense physical activity (Nyborg, 1984). Interestingly, a series of papers has documented the tendency toward either psychological androgyny or masculinity among female athletes (Burke, 1986; del Rey & Sheppard, 1981; Jackson & Marsh, 1986; Myers & Lips, 1978; Pamich, 1987).

A striking feature of these reports is their failure to make a connection between *psychological* and *biological* androgyny. The latter is recognized to result from low levels of feminizing hormones and/or high levels of masculinizing hormones. These are the very influences that offer the most parsimonious explanation for observed differences between HMHM females and controls.

That these hormones influence not only somatic traits but also personality and behavior has been amply documented (Forest, 1983; Nyborg, 1984; Udry, Morris, & Kovenock, 1995). Moreover, while the focus thus far has been on postnatal hormones, evidence suggests that activity levels throughout life also reflect prenatal hormone exposure (Forest, 1983; Hurtig, Radhakrishnan, Reyes, & Rosenthal, 1983; Nyborg, 1984).

As noted earlier, females with congenital adrenal hyperplasia are exposed to high levels of testosterone in utero (Nyborg, 1984). Most cases are diagnosed at birth--with prompt normalization of androgen levels. Nonetheless, affected females retain certain traits as a consequence of a prenatal environment more similar to that of a male fetus. A tendency toward intense physical activity is one such characteristic (Hurtig et al., 1983).

This is not to suggest that HMHM females were exposed in utero to testosterone levels comparable to those seen in CAH. It is nonetheless plausible that differences between HMHM females and controls have been partially mediated by differences in the prenatal hormonal environment.

Intriguing evidence has suggested that prenatal hormone exposure sets the stage for postnatal life. One study found a correlation between exposure to bioavailable testosterone during gestation and the amount secreted by women in adult life (Udry et al., 1995). Moreover, the magnitude of masculine interests, appearance, and behavior in these women correlated with the degree of exposure to testosterone in utero (Udry et al., 1995).

Handedness. The inconclusive outcome of the prediction regarding handedness is not surprising. When the percentage of individuals who possess a given trait is small, significant differences between groups are likely to be detected only when sample sizes are large (Benbow, 1986).

Fewer than 10% of all subjects were left-handed. Moreover, NLSY interviewers did not inquire about handedness until 1992. As a consequence, only 80 of the original 127 subjects provided information about hand preference. This is a small sample given the rarity of the trait.

The assessment technique was also less than optimal, and perhaps inadequate. Subjects were simply asked, Were you born naturally right or left handed? Whether asking about hand preferences at birth rather than at the time of interview created confusion cannot be determined. What can be said is that treating handedness as an either-or trait provides only a crude measure of hand preference (Reiss, Reiss, & Freye, 1998).

Contrary to common belief, hand preference is not a simple choice of right or left. Rather, it is a characteristic that falls along a continuum. Some individuals use the left hand for writing, but favor the right hand for throwing or other activities (Gilbert & Wysocki, 1992). Accordingly, handedness is best measured with an inventory that measures the degree of preference for the right or left hand.

Several such tools are available. The Edinburgh Handedness Inventory is one example. This 10-item questionnaire inquires about preferred hand for a variety of tasks and is widely used in research (Oldfield, 1971). Use of a questionnaire such as this one, which measures handedness in degrees, may have produced a different result. Even were the outcome identical, confidence in the results would nonetheless increase.

Both left-handedness and mixed-handedness in females have been associated with more masculine behaviors and sex-role identification (Casey & Nuttall, 1990). Prenatal exposure to testosterone also has been cited as a factor in determining left-hand preference (Geschwind & Behan, 1982). Because other findings of this study also implicate testosterone, discarding the prediction about handedness would be premature.

Further testing with larger samples and a better assessment tool is needed to reach a conclusion about this aspect of the dissertation model. In light of the interaction between familial handedness and spatial ability, information about handedness of biological relatives should be sought as well (Casey & Brabeck, 1990; Casey et al., 1992a; Casey, Pezaris, & Nuttall, 1992b). Such data were not available for evaluation here.

Finally, handedness and ability should be assessed concurrently. Ability testing of NLSY participants took place in 1980, while handedness was not inquired about until

1992. While little evidence exists that hand preference changes dramatically over time, several reports document shifts in degree of hand preference across the lifespan (Bryden, MacRae, & Steenhuis, 1991; Gilbert & Wysocki, 1992). Simultaneous assessment of handedness and ability is therefore preferable.

Traditionality. As with handedness, results regarding traditionality proved inconclusive. In retrospect, the most obvious problem was the primary measurement tool.

The psychometric measure of sex-role traditionality was the NLSY Scale of Traditional Values (STA). The scale has never been validated; accordingly, its predictive value remains uncertain. Contributing to concerns about its validity is a report that high scores on the scale predicted risk of becoming an unwed teenage mother (Plotnick & Butler, 1991). While childbearing is a traditional role for females, doing so outside of marriage arguably is not.

The scale was probably unsuitable for testing the dissertation model because it involves general statements about *all* women. Statements of this kind may elicit a response that says less about a subject's own values than her views about other women.

For example, one item reads, Women are happier if they stay home and take care of their children. A subject who rejects the housewife role may nonetheless believe that most women prefer it. In such a case, the STA score would likely be a poor measure of traditionality. Regrettably, the NLSY did not include any scale that assessed the sex-role identity of subjects as individuals.

Range restriction may also have contributed to the failure of the scale to distinguish between HMHM females and controls. Even if valid when used in samples representative of the U.S. population, the scale may distinguish poorly among specific subgroups. This study included only females who scored in the top quartile of general ability. Statements such as A woman's place is in the home may fall on uniformly deaf ears among such a select group.

Traditionality was also assessed by responses to a series of questions about work and family aspirations. The timing of these may also have contributed to the mixed pattern of results. These questions were generally posed on entry into the NLSY, when subjects ranged in age from 14 to 22. Conceivably, some were too young at the time to accurately report the desires they would have as adults.

Despite expressing rather traditional aspirations in 1979, HMHM females went on to lead less conventional lives than controls. They were less likely to marry, have children, or work in female-dominated occupations. One could therefore argue that the prediction was affirmed. However, because lifestyle conflicted with stated values and aspirations, the data are better viewed as inconclusive. Further research is justified and should include tools that assess an individual's sex-role identity.

Researchers at the Carolina Population Center have developed a comprehensive approach to measuring sex-role identity in females (Udry et al., 1995). The assessments consider how strongly a subject manifests a feminine identity in her interests, career, mannerisms, and appearance. Use of such an approach is likely to yield more conclusive results.

Tertiary Variables

Religiosity. This is the only tertiary variable that yielded a statistically significant result. As predicted, HMHM females attended religious services less frequently than controls. This is consistent with the lesser religiosity of males and supports an inference that HMHM females are less sexually differentiated than controls. However, a hormonal mechanism that specifically contributes to lesser religiosity has not been elucidated.

Such an effect is not beyond the realm of reason. Hormones clearly influence the organization of the brain, as well as its ongoing functions (Hampson, 1995; VanGoozen, 1994; VanGoozen et al., 1995). Among the characteristics affected are deductive reasoning skills (Kimura & Hampson, 1993). Individuals predisposed to a high degree of rational, deductive thought may have difficulty with religious doctrines that must be accepted as a matter of faith. Nonetheless, no data are available that directly address this hypothesis.

Whether future research should include religiosity is difficult to say. On the one hand, this construct has proved useful in differentiating HMHM females from controls. On the other, its mechanism remains elusive. Accordingly, its importance for testing models rooted in GTC theory cannot be argued with much certainty.

Depressive symptoms. The failure of the prediction regarding depressive symptoms is difficult to attribute to the test instrument. Of the psychological inventories used to assess emotional well-being, the CES depression scale has been particularly well studied. As discussed in chapter 3, the literature confirming its validity is both extensive and diverse.

Admittedly, none of these validation studies specifically examined its use in a population restricted to the higher deciles of general ability. However, average intelligence does vary across the many population groups in which the instrument has been validated. The restricted range of general ability studied here is therefore unlikely to explain the insignificant difference in depressive symptoms between HMHM and control females.

Several alternative explanations merit consideration. The first is simply the small number of subjects. NLSY interviewers did not administer the CES-D scale until 1992. As a result, fewer subjects completed the scale than would have been the case with earlier administration of the inventory.

In addition, both constitutional and situational factors contribute to depressive symptoms (Svrakic et al., 1991). In this regard, the difference in marriage rates between HMHM and control females warrants consideration. As of 1992, HMHM females were considerably more likely than controls to have been single all their lives.

Reexamination of the depression data within each ability group revealed a trend toward higher scores among those who had never married. The association was particularly striking within the HMHM group. The mean score of HMHM females who had never married exceeded that of other group members by more than three points.

In the context of this study, a three-point difference is substantial. After all, the mean CES-D score of all control females exceeded that of all HMHM females by less than a half point. This suggests that control for situational factors may be needed. The NLSY data do not lend well to such an approach, as most of the life history data are not from the same year as the CES-D score.

Differential rates of depression between males and females emerge around the time of puberty (Angold & Worthman, 1993). This temporal relationship suggests an influence of hormones and supports retention of this variable in the study model. However, future testing will require a larger sample than available here. It would also be advisable to favor young adults when selecting subjects. Such an approach may minimize the potential for confounding by situational factors related to attainment of adult goals.

Locus of control and self-esteem. The study model anticipated that HMHM females would manifest higher self-esteem and a more internal locus of control. These predictions were rooted in earlier reports that males manifest these characteristics more strongly than females. Analysis of scoring patterns for all males and females in the NLSY provided a ready explanation for the failure of these predictions. The test instruments used here produced no *meaningful* sex differences in scores when administered to this large, carefully sampled population.

NLSY subjects completed the Rosenberg Self-Esteem scale in 1980 and again in 1987. On both occasions, males scored only fractionally higher than females. The difference in mean score between sexes was 0.3 points on a 40-point scale. For the Pearlin Mastery Scale, the difference between sexes again favored males. However, it was extremely small: 0.2 on a scale with a maximum value of 28.

These findings highlight the importance of distinguishing between statistical and clinical significance, particularly when working with large samples. When nine to twelve thousand people are tested, the statistical significance of minute differences may be high. The self-esteem inventory is a case in point; the mean difference of 0.3 points between NLSY males and NLSY females was highly significant ($p < .001$).

From a practical standpoint, it is difficult to argue that these differences actually merit the term. Evidence is lacking that such a small difference in scores would predict disparate life outcomes.

Alternatively, the lack of meaningful differences in test scores between males and females could indicate a problem with the instrument itself. However, this seems doubtful. Feingold conducted a meta-analysis of dozens of studies and concluded that any disparity in self-esteem between the sexes is very small (Feingold, 1994). This is consistent with the negligible difference in mean RSE scores among NLSY males and females.

Feingold also analyzed numerous studies that used self-report inventories to assess internality. He found no overall differences between the sexes. This underscores that the failure of the prediction regarding internality should not be attributed to the test instrument, but to the model.

Simply stated, internality and self-esteem are not useful constructs for testing conceptual models based in GTC theory. Both are best excluded from future research. Characteristics that show more meaningful differences between the sexes may be substituted instead.

The most suitable alternatives would be tender-mindedness (also called nurturance) and assertiveness. Females manifest far more tender-mindedness than do males (Feingold, 1994). By contrast, males are more assertive. These differences are sufficiently well-established to be incorporated into future theoretical models.

Recent evidence of an inverse relationship between free testosterone and nurturance--both within and between the sexes--further supports such a change (Harris, Rushton, Hampson, & Jackson, 1996). The same study also found testosterone to be correlated with self-reported aggressive traits. To the extent that assertiveness overlaps with aggressiveness, this finding also argues for modifying the study model.

Anxiety is another trait that shows a marked sex difference, with females more commonly affected (Geary, 1998a). However, it often occurs in tandem with depressive symptoms. This cluster of traits is already represented in the model--as well as on the CES-D scale (Orme, Reis, & Herz, 1986). The value of incorporating it as an additional variable is therefore debatable.

With these suggested modifications, a revised model would predict HMHM females to be more assertive, but less tenderminded, than other females. The NLSY database contains no information suitable for evaluating these predictions.

Quality of Data Considerations

In any investigation, data quality is a paramount concern. The NLSY is widely recognized in the social sciences. Nonetheless, it was designed to assess labor force behavior--a purpose very different from the use for which it was employed here. Whether its data were adequate for testing the predictions of this study is a reasonable question.

It is doubtful that any variable in the study is immune from criticism of some kind. Height and weight, for example, were not based on measurements taken by interviewers, but on self-reports. Contraceptive use was framed in terms of the month prior to interview--a measure not necessarily representative of behavior for a full year. Data are insufficient to allow for consideration of marital duration as a co-variate in coital frequency--a factor that may be relevant (Udry, 1980). Certain data--such as menarcheal age--require recall of events that occurred as much as a decade earlier. The list of potential objections to the data set may well be endless.

Certain operational definitions are also vulnerable to criticism. Use of church attendance to operationalize religiosity can be questioned; those who attend regularly may do so at the behest of a spouse or other consideration unrelated to faith. Use of coital frequency to operationalize libido is similarly open to challenge. Few would deny that libido encompasses a range of thoughts and behaviors beyond intercourse alone.

In response, one can cite evidence that all operational definitions have precedents in the relevant literature, that self-reported data are often sufficiently accurate, and that some instruments used in the study have substantial validity and reliability. Moreover, no reason exists to believe that errors of self-reporting would bias the data one way in the HMHM group and in the opposite direction among the controls.

Nonetheless, a single question rises above all such concerns. It is whether the consistent findings reported here can be dismissed as an artifact arising out of deficiencies in the data or the operational definitions.

The answer is a straightforward no. The number of findings here consistent with GTC theory is too great to be dismissed based on concerns about a particular measurement tool or definition. Moreover, the distribution of significant findings strengthens the case considerably.

Significant results were shown for all primary variables. Of seven secondary variables, predictions were supported for five and deemed inconclusive for two. While three predictions were rejected, all pertained to tertiary variables--one of which was affirmed.

This is not to say that these data are conclusive, for they are not. It is simply to acknowledge the evidence here as sufficient to support a new approach to the study of sex differences--one that recognizes the role of hormones in developmental processes.

The Outcomes in Perspective

This study supports a variety of conclusions--some definitive, some tentative. The most definitive statement supported by the findings is that HMHM females are biologically different from controls. One can also say with confidence that HMHM females are less sexually differentiated. Their life histories are unusual--and in all likelihood, mediated by hormone profiles uncharacteristic of most females.

The conclusions that can be drawn about the nature of these hormonal differences are more tentative. Certainly, lesser exposure to estrogens is implicated very strongly. Greater exposure to androgens during key periods of gestation and higher androgen levels in adult life are also compelling candidates. Whether these are independent of estrogen status remains unclear. Determination of specific hormonal influences will require further research, using more precise tools and longitudinal evaluation of endocrine status.

The question that remains is whether the lesser sexual differentiation of HMHM females is merely a marker for their unusual ability pattern, or a cause of it. The findings here have demonstrated associations between the HMHM pattern and certain somatic and reproductive characteristics. Can the case be made that all are dependent variables with a common causality?

The answer is a qualified yes. Chapter 2 reviewed an impressive array of literature that links estrogens, free testosterone, or both to cognitive pattern. This research yielded a pattern of results for which the sole parsimonious explanation is the hormonal milieu (or an extraordinary level of coincidence). Moreover, the tendency for HIIQ females to fall in a mid-range between the two groups on an assortment of traits is consistent with exposure-related effects.

Of the three ability groups studied here, HMHM females appear to represent the most extreme manifestation of the causative factor(s). HIIQ females are intermediate, and controls the most modest. The causative factor *may* be the relative balance of free testosterone to estradiol during the second trimester of pregnancy and other periods of postnatal life and/or the relative degree of exposure to key estrogenic compounds. This cannot be determined from the kind of data available here, nor from the current state of knowledge.

One must also recognize that hormones have both direct and subtle impacts (Nyborg, 1994). The hormonal milieu affects more than cognitive structures (Geary, 1996). It also influences how individuals experience and interact with their environments. These effects have profound implications for development of particular faculties (Purifoy, 1981).

This concept is particularly important in regard to mechanical comprehension. Because this subject is rarely taught in school, HMHM females must have had more than a natural tendency to excel in this area. They must also have had the interest to learn about mechanical devices and principles on their own. Doing so required that they select, from the many activities available to them, those that facilitated development of their mechanical talents. What motivates females to tinker with mechanical objects rather than choose more sex-typical activities is an important part of this story. It serves to remind that the word cause is best used with care.

One final conclusion requires mention. It is that unusual cognitive abilities in women can no longer be considered independent of biological influences. Continued efforts to do so are simply spurious.

For example, numerous reports have linked high mathematical ability in females with an androgynous or masculine score on sex-role inventories (Signorella & Jamison, 1986). For decades, psychologists have attributed sex-role development to internalization of the characteristics of the same-sex parent (Juni, Rahamim, & Brannon, 1985). Accordingly, a girl with a masculine self-concept was often said to have identified with her father.

In the context of life history data, such a perspective is no longer sustainable. Perhaps it is reasonable to suspect that a girl who prefers trains to dolls or pants to dresses identifies with her father. But to suggest that such a child can somehow will herself to be taller, thinner, older at menarche, and at high risk of losing a pregnancy is beyond the realm of reason. These characteristics simply cannot be explained using the amorphous notion of a masculine role identity.

Moreover, if a masculine identity were causally related to the cognitive abilities in question, the most mathematically and spatially gifted males would be highly sex-typed. An impressive array of evidence shows otherwise. Males who excel in these domains are often not very masculine at all.

As measured by psychological inventories, students with outstanding reasoning or spatial abilities tend to be the least sex-typed: masculine females and feminine males (Lippa, 1998a; Lubinski & Humphreys, 1990; Signorella & Jamison, 1986). These outcomes are reminiscent of findings regarding salivary testosterone and spatial competence. The more able females show high levels relative to other women, while in males the pattern is reversed (Gouchie & Kimura, 1991).

This is compelling evidence that high mathematical-spatial skill depends on an optimal balance of hormones possessed by the more masculine females and the more feminine males. More sex-typed members of both sexes apparently fall outside this optimum range (Nyborg, 1994).

Once again, it must be stressed that the magnitude of sexual differentiation is not the sole influence on the variables studied here. It is, however, the common thread--one that ties the HMHM ability pattern in females and numerous variables of the dissertation model to a shared hormonal milieu.

A Case of Sexual Selection

This investigation had two goals. The first was to test a set of predictions about the correlates of the HMHM ability pattern. The second was to gain insight into the rarity of this cognitive pattern in females.

Few would deny that sex differences in cognition have received a large--if not inordinate--amount of attention. Yet, the rarity of the HMHM pattern in females remains poorly understood. That a better understanding of its origins has eluded psychologists raises an important question. It is whether the jury has remained out by virtue of its failure to consider the most crucial evidence.

In an insightful commentary about the mindset of contemporary psychology, Silverman and Phillips observed:

> Scientists deal with diverse levels of causation, which can be conceptualized on a continuum between proximate and ultimate. Proximate refers to more immediate causes--ultimate to those more remote . . . .
>
> The theory of evolution by natural selection remains the sole scientific theory of the ultimate origins of animal and human behavior. Thus, by its virtual eschewal of this theory, traditional psychology has remained wholly proximate in its paradigms. Psychologists' attempts to understand the processes that propel and direct behavior are limited to *ontogenetic* development--events that occur during the life span of an organism. *Phylogenetic* development, comprising evolutionary origins of behavior has an insignificant role in traditional psychological theorizing. (Silverman & Phillips, 1998, pp. 604-605)

The nascent science of evolutionary psychology seeks to close this theoretical gap by considering both ultimate and proximate causes of behavior. This emerging discipline studies the psyche in light of current knowledge and theory about the evolutionary processes that created it (Daly & Wilson, 1983, p. 254). Unlike conventional approaches, evolutionary psychology sees cognition and behavior as characteristics that have been shaped by the same selection pressures as human morphology.

These selection pressures date back to the Pleistocene--an era that began about three million years before present and ended only 10,000 years ago (Diamond, 1998). During this time, man evolved from ancestral forms into modern-day *Homo sapiens.* Evolutionary psychologists believe that traits common among humans today somehow enhanced fitness in the prehistoric environment. It is often referred to as the EEA--the Environment of Evolutionary Adaptedness (Tooby & Cosmides, 1989). Conceptually, the EEA should be understood not as a specific place or time, but as a composite of the many problems faced by humans in ancestral environments.

For humans in the EEA, one aspect of life was invariant. Without exception, all lived as members of hunting and gathering groups (Daly & Wilson, 1983). Archeological research, as well as studies conducted among present-day hunter-gatherers, has provided valuable information about the realities of these environments. These permit insight into the kind of traits that might have enhanced survival in the EEA.

Some problems facing humans in the EEA differed little--if at all--between the sexes (Buss, 1995). Morphological similarities between males and females reflect this reality. For example, both sexes have sweat glands, the shivering response, and a host of other characteristics that reflect their common need to regulate body temperature. Of course, every trait has a price, or trade-off. However, unless the disadvantages it confers outweigh advantages, natural selection favors its retention (Darwin, 1859).

Also important to evolutionary theory is a process Darwin called sexual selection. He mentioned it but briefly in his classic work, *On the Origins of Species by Means of Natural Selection.* However, in a later book, *The Descent of Man and Selection in Relation to Sex,* Darwin developed the concept more fully.

In Darwin's words, sexual selection is a process that arises from the advantage which certain individuals have over other individuals of the same sex and species, in exclusive relation to reproduction (Darwin, 1871, p. 256). Simply stated, sexual selection favors traits conferring individuals with a *reproductive* advantage over their same-sex competitors. Such characteristics are those that enhance an individual's ability to attract mates and successfully rear offspring resulting from the union.

Evolutionary psychologists have begun to consider cognitive sex differences in the context of sexual selection. In the process, a viable explanation for the greater spatial ability of males has emerged. (The concept that mechanical ability is comprised largely of spatial skill, as well as evidence that spatial ability enhances mathematical competence, was reviewed earlier and will not be repeated here.)

One element of this explanation relates to the very ability to find mates. In humans, securing a mate is more difficult for males than females (Daly & Wilson, 1983). As a consequence, competition between males for willing partners can be considerable--if not fierce.

In the EEA, the ability to navigate novel or difficult terrain may have enhanced the likelihood of males finding females for mating and reproduction (Geary, 1998a). If so, selection pressures for the retention of spatial ability would have been stronger for males. This is consistent with findings from a wide range of species. These show that differences in spatial ability predictably favor the sex whose reproductive strategy requires better skill at negotiating the physical environment (Geary, 1998a).

Another element of this thinking focuses on the relationship between spatial ability and hunting prowess. A sexual division of labor in which males hunted and females foraged has been a distinguishing feature of hunter-gatherer societies across time (Silverman & Eals, 1992). Silverman and Eals argued that the requisite cognitive skills differ:

> Tracking and killing animals entail different kinds of spatial problems . . . The cognitive mechanisms of contemporary *Homo sapiens* appear to reflect these differences, insofar as the various measures showing male bias (e.g. mental rotations, map reading, maze learning) correspond to attributes that would enable successful hunting. Essentially these attributes comprise the abilities to orient oneself in relation to objects or places, in view or conceptualized across distances, and to perform the mental transformations necessary to maintain accurate orientation during movement. This would enable the pursuit of prey animals across unfamiliar territory and, also, accurate placement of projectiles to kill or stun the quarry. (pp. 534-535)

Geary has expanded this concept to include its psychosocial sequelae. His integrated model posits that sexual selection has acted via biological influences (e.g. hormones) to effect sex differences not only in navigational skills, but also social styles and preferences (Geary, 1998b). The latter contribute to sex differences in perceptions about the usefulness of mathematics--with resultant disparities in course-taking and participation in activities that support development of innate potentials. Coupled with the advantage conferred by certain superior spatial abilities, these psychosocial factors exacerbate sex differences in the cognitive domains at issue here.

This is not to say the superior spatial ability of males has come at no cost. As noted earlier, evolutionary theory anticipates that each trait carries a price. Findings from this study suggested a trade-off between technical and verbal abilities. Similarly, an inverse relationship between fluent production and spatial ability in adolescent males has been reported (Petersen, 1976). However, for purposes of this discussion, it is net effect that matters most. Available evidence suggests that whatever its cost, the net value to males of high spatial ability is positive.

Cross-cultural research has confirmed that females--including those in present-day hunter-gatherer societies--prefer mates who demonstrate provisioning ability (Buss, 1995; Daly & Wilson, 1983; Geary, 1998a). This preference makes sense; a male's provisioning ability has obvious implications for the survival chances of his mate(s) and their offspring. Success in hunting is a classic demonstration of provisioning ability (Shepard, 1998).

One might even argue that female preference for males with a high spatial-mechanical skill was evident in the NLSY. Compared to their HIIQ counterparts, HMHM males married and fathered children at a younger age. How much meaning--if any--should be attached to this disparity is debatable. The critical finding is that HMHM males suffered no consequences relative to marriage or reproduction. The same cannot be said for HMHM females--and is perhaps the central finding of this study.

The life histories of HMHM females add a new dimension to the theoretical framework presented by Silverman and Eals and by Geary. In their conceptual models, sexual selection worked actively to retain spatial competence in males; the poorer skills of females emerged as a consequence. The data here are consistent with the notion that selection pressures favored the preservation and enhancement of these abilities in males. However, the findings also suggest that sexual selection worked actively, rather than passively, in females--with lesser spatial ability one of its consequences.

Sexual selection is a bottom-line process. That which enhances a female's reproductive success relative to other women is made commonplace; that which does not is rendered rare. Nonetheless, the process need not act directly on every characteristic of living organisms. Rather, a trait may be selectively retained--or conversely, rejected--simply by virtue of its association with characteristics that enhance or detract from fitness. The evolutionary fate of the cognitive structures and processes that underlie the HMHM pattern was probably determined not by their inherent value, but the constellation of characteristics that accompanied them.

Assuming that the results of this study are replicable, the rarity of the HMHM pattern in females is readily demystified. From an evolutionary standpoint, the value of the pattern is dwarfed by its attendant costs. Females with cognitive structures and processes conducive to the HMHM pattern may have had any number of traits useful in the EEA. It is doubtful that a reproductive advantage over other females was one of them.

A number of characteristics associated with the HMHM pattern illustrate this point. For this discussion, only two need mention. The first is the extraordinary rate of pregnancy loss in HMHM females. Despite keen intelligence, absence of known risk factors, and residence in a country with abundant food and medical care, these females lost almost one in four of their pregnancies.

Conditions in the evolutionary environment--or in pre-industrial Europe, for that matter--were not nearly as favorable. Rather, food shortages and seasonal access to foods containing nutrients critical to fetal growth were common (Frisch, 1990; Knapp, 1998). Antibiotics and modern medical technology were absent. One cannot help but speculate that in such environments, the pregnancy losses of HMHM females would have been higher still.

Body fatness needs to be viewed in evolutionary perspective as well. That HMHM females were leaner than controls is an important consideration. Also critical was the finding that HMHM females who never reproduced averaged a lower BMI in 1981 than group members who eventually bore children.

In environments marked not only by famine and seasonal food shortages, but also mandatory lactation, body fat was more crucial to reproductive success than it is today (Frisch, 1990). Although lactation is now optional in industrial societies, nursing women are advised to consume an extra 500 calories per day (FNB, 1990). This allotment increases by 30% if a woman does not gain sufficient weight during pregnancy or becomes underweight while breast-feeding.

Assuming the lesser figure, this recommendation translates into a pound of stored fat for every week that a nursing woman cannot obtain additional calories from the available food supply. Were such a situation to exist for 3 months, a woman would need 12 pounds of stored fat that could be sacrificed for milk production. Given the very long periods of nursing characteristic of the non-industrialized world (Eaton et al., 1994), significantly greater amounts of stored fat may well have been important.

In addition to enhancing her reproductive potential, good fat reserves may have been critical to a woman's own survival. Prior to effective control of communicable diseases, stored fat was valuable for weathering infection (Weigley, 1984). Inadequate energy intake combined with rigorous physical activity is now known to suppress immunocompetence (Kramer et al., 1997). During periods of food restriction, those with few fat reserves may have succumbed more readily to death from infectious disease. If such a death occurred prior to achieving full childbearing potential, a woman's reproductive success would have been compromised.

Figure 19 summarizes hypothesized mechanisms by which a high androgen to estrogen ratio may have inhibited female reproductive success in ancestral environments. If accurate, sexual selection would have operated to limit the occurrence of this profile in females. In the process, its correlates would have become increasingly rare. This appears to be the most promising explanation for the rarity of the HMHM ability pattern in females.

Testing of hypotheses based in evolutionary theory is naturally difficult. Researchers cannot recreate the ancestral environment--at least not in the developed world. Owing to widespread industrialization, no European population retains a lifestyle reminiscent of evolutionary conditions. However, several hunter-gatherer societies in Africa do provide the opportunity to study humans in settings considered reasonable approximations to the EEA. These investigations have provided support for some of the concepts elucidated here.

Among African hunter-gathers, the best-studied population is the !Kung. These sub-Saharan foragers live in conditions considered most representative of those present during the Pleistocene era (Blurton Jones, Smith, O'Connell, Hawkes, & Kamuzora, 1992). Investigations conducted among a lesser-known tribe of hunter-gathers--the Hadza of Tanzania--also provide insights into the correlates of reproductive success in an environment reminiscent of the EEA.

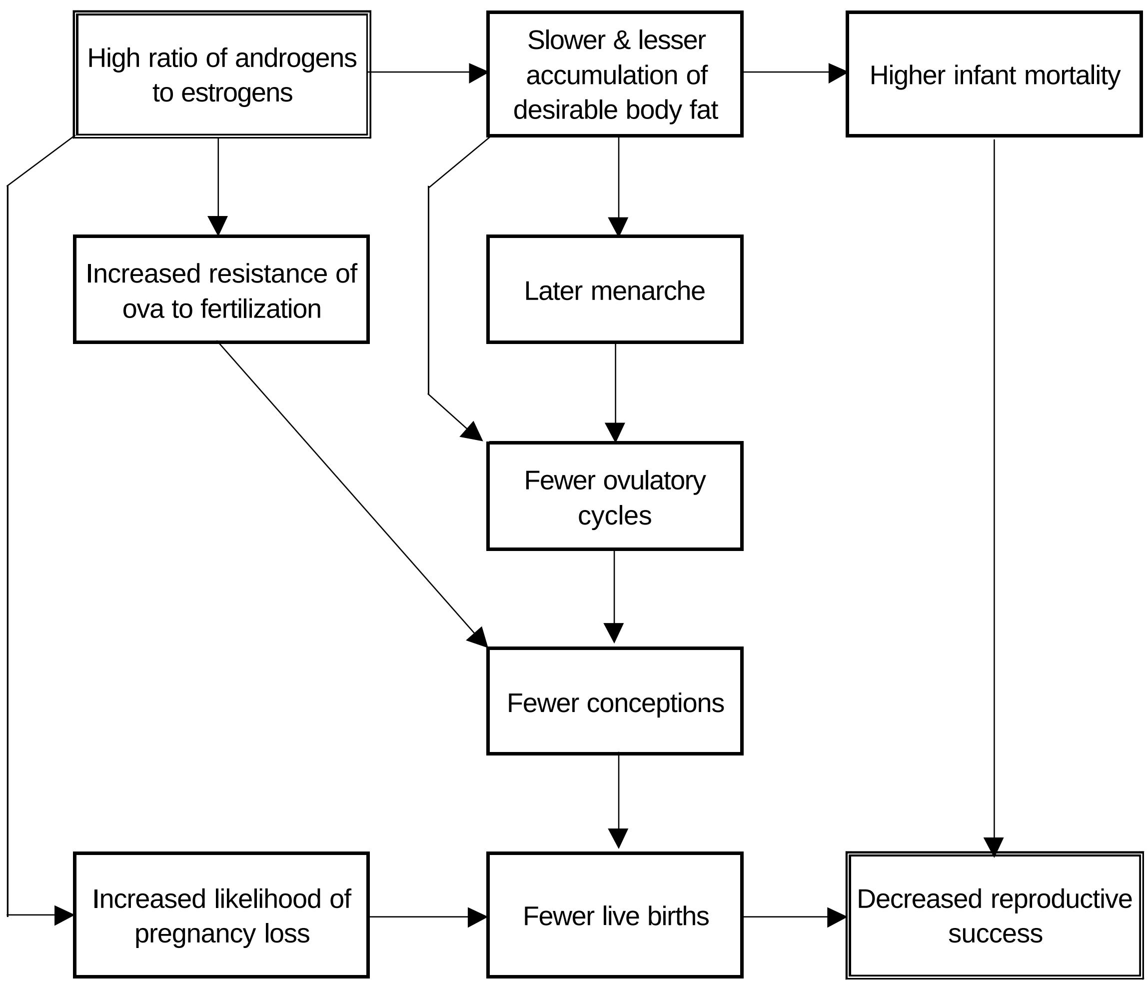


Figure 19. Hormone balance and reproductive success.

Hypothesized mechanisms by which a high ratio of androgens to estrogens inhibited the reproductive success of females in ancestral environments.

Interestingly, Hadza women have greater reproductive success than their !Kung counterparts (Blurton Jones et al., 1992). The average number of offspring is 6.2 for Hadza females, and 4.7 for the !Kung. Moreover, during their 30s, Hadza females have higher fertility than !Kung women of comparable age. Indeed, Hadza females show high fertility after the age of 35. It is not uncommon for them to give birth after the age of 40.

Body composition may well explain the differential outcomes. Though similar in height, Hadza women are heavier than their !Kung counterparts by an average 17 pounds (Hiernaux & Boedhi Hartono, 1980). In fact, the substantial subcutaneous fat of Hadza women is considered a characteristic feature that distinguishes them from females of other African societies.

Their continued fertility at a late age is particularly notable. In industrialized societies, fertility decreases by an estimated 50% after 40, and the risk of pregnancy loss increases by 100% or more (Toner & Flood, 1993). The substantial body fat of Hadza women may explain their fertility well into the fifth decade. Estrogens produced by their fat reserves may create a hormonal environment conducive to successful conception and pregnancy.

The traditions of the Kipsigis also deserve mention. This Kenyan tribe is not as removed from all modern influences as are classic hunter-gatherers. Nonetheless, it is a largely agrarian society where contraception is not practiced (Borgerhoff-Mulder, 1989b). Accordingly, it offers some evidence regarding the historical correlates of reproductive success.

The relationship between later menarche and reduced reproductive success (Table 19) is well understood by these indigenous people. Kipsigis males bid for a wife by offering items of value to her father. The amount of the offering varies with her menarcheal age. Early-maturing females attract higher bids than late-maturers (Borgerhoff-Mulder, 1989a). Demographic research confirms that the early maturers have longer reproductive life spans and better fertility (Borgerhoff-Mulder, 1989b).

To summarize, sexual selection has shaped the cognitive, behavioral, and morphological characteristics of males and females. During the course of hominid evolution, it has selected against traits associated with reduced reproductive success. In females, characteristics that underlie the HMHM ability pattern appear to fall into this category.

In other words, the rarity of this cognitive pattern in females is not enigmatic. To the contrary, it is a predictable outcome of the process that has shaped all forms of life on Earth.

Conclusion

Analysis of data from the National Longitudinal Survey of Youth revealed that females with high levels of mathematical and mechanical ability have life histories that differ significantly from those of other academically able women. Somatic and reproductive characteristics of these subjects suggest that their magnitude of sexual differentiation is very low. Lesser exposure to estrogens and/or a high ratio of bioavailable testosterone to estrogen may well be causative factors.

The findings strongly support concepts underlying Helmuth Nyborg's theory of general trait co-variance. Further, the results indicate that the rarity of the HMHM ability pattern in females derives primarily from biological phenomena.

Admittedly, these conclusions challenge longstanding notions about the origins of sex differences in ability and temperament. This is not unusual, but consistent with the natural process of scholarly inquiry. Further investigation will determine whether the interpretations offered here are ultimately accepted, rejected, or modified.

The most immediate need is to replicate the outcomes here using larger populations and more exacting methods. Research that examines fertility among HMHM females who attempt pregnancy is especially needed. Also important is more precise knowledge about the physiological mechanisms that influence cognitive skills. Finally, an understanding must be developed regarding the trade-offs that invariably would attend any effort to modify cognitive outcomes in females.

Research into the biological aspects of cognitive sex differences serves broad scientific and societal goals. An understanding of the etiology of cognitive processes should help psychologists and educators better serve males and females alike--as well as allow the public policy issues raised by cognitive sex differences to be put into a rational context.

References

AAUW. (1989). Equitable treatment of girls and boys in the classroom. Washington, D.C.: American Association of University Women.

AAUW. (1991). Shortchanging girls, shortchanging America. Washington, D.C.: American Association of University Women.

AAUWEF (Ed.). (1998). *Separated by sex: A critical look at single-sex education for girls*. Washington, D.C.: American Association of University Women Educational Foundation.

Abrams, B., Carmichael, S., & Selvin, S. (1995). Factors associated with the pattern of maternal weight gain during pregnancy. *Obstetrics and Gynecology, 86*(2), 170-176.

Ainsworth, B. E., Sternfeld, B., Slattery, M. L., Daguise, V., & Zahm, S. H. (1998). Physical activity and breast cancer: Evaluation of physical activity assessment methods. *Cancer, 83*(3 Suppl), 611-620.

Aksoy, S., Celikkanat, H., Senoz, S., & Gokmen, O. (1996). The prognostic value of serum estradiol, progesterone, testosterone and free testosterone levels in detecting early abortions. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 67*(1), 5-8.

Albanes, D., & Taylor, P. R. (1990). International differences in body height and weight and their relationship to cancer incidence. *Nutrition and Cancer, 14*(1), 69-77.

Allison, D. B., Kaprio, J., Korkeila, M., Koskenvuo, M., Neale, M. C., & Hayakawa, K. (1996). The heritability of body mass index among an international sample of monozygotic twins reared apart. *International Journal of Obesity and Related Metabolic Disorders, 20*(6), 501-506.

Altermatt, E. R., Jovanovic, J., & Perry, M. (1998). Bias or responsivity? Sex and achievement-level effects on teachers' classroom questioning practices. *Journal of Educational Psychology*, *90*(3), 516-527.

Andersson-Ellstrom, A., Forssman, L., & Milsom, I. (1996). Age of sexual debut related to life-style and reproductive health factors in a group of Swedish teenage girls. *Acta Obstetricia et Gynecologia Scandinavica, 75*(5), 484-489.

Andreason, P. J., Zametkin, A. J., Guo, A. C., Baldwin, P., & Cohen, R. M. (1994). Gender-related differences in regional cerebral glucose metabolism in normal volunteers. *Psychiatry Research, 51*(2), 175-183.

Aneshensel, C. S., Pearlin, L. I., & Schuler, R. H. (1993). Stress, role captivity, and the cessation of caregiving. *Journal of Health and Social Behavior, 34*(1), 54-70.

Angold, A., & Worthman, C. W. (1993). Puberty onset of gender differences in rates of depression: A developmental, epidemiologic and neuroendocrine perspective. *Journal of Affective Disorders, 29*(2-3), 145-158.

Annett, M. (1979). Family handedness in three generations predicted by the right shift theory. *Annals of Human Genetics, 42*(4), 479-491.

Annett, M. (1992). Spatial ability in subgroups of left- and right-handers. *British Journal of Psychology, 83*(4), 493-515.

Anonymous. (1998). *The Rosenberg Self-esteem Scale.* [On-line]. Available: http://www.bsos.umd.edu/socy/rosenberg.htm.

Apter, D., Reinila, M., & Vihko, R. (1989). Some endocrine characteristics of early menarche, a risk factor for breast cancer, are preserved into adulthood. *International Journal of Cancer, 44*(5), 783-787.

Apter, D., & Vihko, R. (1985). Premenarcheal endocrine changes in relation to age at menarche. *Clinical Endocrinology, 22*(6), 753-760.

Apter, D., & Vihko, R. (1990). Endocrine determinants of fertility: Serum androgen concentrations during follow-up of adolescents into the third decade of life. *Journal of Clinical Endocrinology and Metabolism, 71*(4), 970-974.

Arushanyan, E. B., & Borovkova, G. K. (1989). Psychotropic properties of ovarian estrogens. *Neuroscience and Behavioral Physiology, 19*(1), 57-66.

Azziz, R., Rittmaster, R. S., Fox, L. M., Bradley, E. L., Jr., Potter, H. D., & Boots, L. R. (1998). Role of the ovary in the adrenal androgen excess of hyperandrogenic women. *Fertility and Sterility, 69*(5), 851-9.

Bagley, C., Bolitho, F., & Bertrand, L. (1997). Norms and construct validity of the Rosenberg self-esteem scale in Canadian high school populations: Implications for counselling. *Canadian Journal of Counselling, 31*(1), 82-92.

Baker, P. C., Keck, C. K., Mott, F. L., & Quinlan, S. V. (1993). *NLSY child handbook*. Columbus, OH: Center for Human Resource Research.

Bakewell, J. M., Stockbauer, J. W., & Schramm, W. F. (1997). Factors associated with repetition of low birthweight: Missouri longitudinal study. *Paediatric and Perinatal Epidemiology, 11 Suppl 1*, 119-1129.

Balakrishnan, T. R., & Chen, J. (1990). Religiosity, nuptiality and reproduction in Canada. *Canadian Review of Sociology and Anthropology, 27*(3), 316-340.

Barr, G. A., Gibbons, J. L., & Moyer, K. E. (1976). Male-female differences and the influence of neonatal and adult testosterone on intraspecies aggression in rats. *Journal of Comparative and Physiological Psychology, 90*(12), 1169-1183.

Barth, R. P., Schinke, S. P., & Maxwell, J. S. (1983). Psychological correlates of teenage motherhood. *Journal of Youth and Adolescence, 12*(6), 471-487.

Baucom, D. H., Besch, P. K., & Callahan, S. (1985). Relation between testosterone concentration, sex role identity, and personality among females. *Journal of Personality and Social Psychology, 48*(5), 1218-1226.

Beatty, W. W. (1984). Hormonal organization of sex differences in play fighting and spatial behavior. *Progress in Brain Research, 61*(7), 315-329.

Becker, J. R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education, 12*(1), 40-53.

Bell, J. (1989). A comparison of science performance and uptake by fifteen-year old boys and girls in co-educational and single-sex schools. APU survey findings. *Educational Studies, 15*(2), 193-203.

Belz, H. F., & Geary, D. C. (1984). Father's occupation and social background: Relation to SAT scores. *American Educational Research Journal, 21*(2), 473-478.

Benbow, C. P. (1986). Physiological correlates of extreme intellectual precocity. *Neuropsychologia, 24*(5), 719-725.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences, 11*(2), 169-183.

Benbow, C. P., & Lubinski, D. (1993). Psychological profiles of the mathematically talented: Some sex differences and evidence supporting their biological basis. *Ciba Foundation Symposium, 178*, 44-59; discussion 60-66.

Benbow, C. P., Zonderman, A. B., & Stanley, J. C. (1983). Assortative marriage and the familiality of cognitive abilities in families of extremely gifted students. *Intelligence, 7*(2), 153-161.

Berkow, R. (Ed.). (1992). *The Merck manual of diagnosis and therapy* (16th ed.). Rahway, NJ: Merck Research Laboratories.

Bernasconi, D., Del Monte, P., Meozzi, M., Randazzo, M., Marugo, A., Badaracco, B., & Marugo, M. (1996). The impact of obesity on hormonal parameters in hirsute and nonhirsute women. *Metabolism, 45*(1), 72-75.

Bertrand, E., Van den Bergh, M., & Englert, Y. (1995). Does zona pellucida thickness influence the fertilization rate? *Human Reproduction, 10*(5), 1189-1193.

Bertrand, E., Van den Bergh, M., & Englert, Y. (1996). Clinical parameters influencing human zona pellucida thickness. *Fertility and Sterility, 66*(3), 408-411.

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119*(3), 422-447.

Binder, G., Grauer, M. L., Wehner, A. V., Wehner, F., & Ranke, M. B. (1997). Outcome in tall stature. Final height and psychological aspects in 220 patients with and without treatment. *European Journal of Pediatrics, 156*(12), 905-910.

BLS. (1998). *National longitudinal surveys*. Bureau of Labor Statistics. [On-line.] Available: http://stats.bls.gov/nlsyouth.htm.

Blurton Jones, N. G., Smith, L. C., O'Connell, J. F., Hawkes, K., & Kamuzora, C. L. (1992). Demography of the Hadza, an increasing and high density population of Savanna foragers. *American Journal of Physical Anthropology, 89*(2), 159-181.

Bock, R. D., & Mislevy, R. J. (1981). Data quality analysis of the Armed Services Vocational Aptitude Battery . Chicago: National Opinion Research Center.

Bock, R. D., & Moore, E. G. J. (1986). *Advantage and disadvantage: A profile of American youth*. Hillsdale, NJ: Erlbaum.

Boklage, C. E. (1981). On the distribution of non-righthandedness among twins and their families. *Acta Geneticae Medicae et Gemellologiae, 30*(3), 167-187.

Borgerhoff-Mulder, M. (1989a). Early maturing Kipsigis women have higher reproductive success than late maturing women and cost more to marry. *Behavioral Ecology and Sociobiology*, *24*(3), 145-153.

Borgerhoff-Mulder, M. (1989b). Menarche, menopause and reproduction in the Kipsigis of Kenya. *Journal of Biosocial Science, 21*(2), 179-192.

Botero-Ruiz, W., Laufer, N., DeCherney, A. H., Polan, M. L., Haseltine, F. P., & Behrman, H. R. (1984). The relationship between follicular fluid steroid concentration and successful fertilization of human oocytes in vitro. *Fertility and Sterility, 41*(6), 820-826.

Boue, J., Bou, A., & Lazar, P. (1975). Retrospective and prospective epidemiological studies of 1500 karyotyped spontaneous human abortions. *Teratology, 12*(1), 11-26.

Bowen, R., South, M., Fischer, D., & Looman, T. (1994). Depression, mastery and number of group sessions attended predict outcome of patients with panic and agoraphobia in a behavioral/medication program. *Canadian Journal of Psychiatry, 39*(5), 283-288.

Bratko, D. (1996). Twin study of verbal and spatial abilities. *Personality and Individual Differences, 21*(7), 621-624.

Brodnick, R. J., & Ree, M. J. (1995). A structural model of academic performance, socioeconomic status, and Spearman's g. *Educational and Psychological Measurement, 55*(4), 583-594.

Brody, L. E., & Benbow, C. P. (1986). Social and emotional adjustment of adolescents extremely talented in verbal or mathematical reasoning. *Journal of Youth and Adolescence, 15*(1), 1-18.

Bronstein, P. M., Wolkoff, F. D., & Levine, W. J. (1975). Sex-related differences in rats open-field activity. *Behavioral Biology, 13*(1), 133-138.

Bruning, P. F., Bonfrer, J. M., Hart, A. A., van Noord, P. A., van der Hoeven, H., Collette, H. J., Battermann, J. J., de Jong-Bakker, M., Nooijen, W. J., & de Waard, F. (1992). Body measurements, estrogen availability and the risk of human breast cancer: A case-control study. *International Journal of Cancer, 51*(1), 14-19.

Bryden, M. P., MacRae, L., & Steenhuis, R. E. (1991). Hand preference in school children. *Developmental Neuropsychology*, *7*(4), 477-486.

Bryden, M. P., & Vrbamcoc, M. L. (1988). Dichotic listening, cognitive ability, and age at puberty. *Developmental Neuropsychology, 4*(2), 169-180.

Burke, K. L. (1986). Comparison of psychological androgyny within a sample of female college athletes who participate in sports traditionally appropriate and traditionally inappropriate for competition by females. *Perceptual and Motor Skills*, *63*(2), 779-782.

Burnett, S. A., Lane, D. M., & Dratt, L. M. (1979). Spatial visualization and sex differences in quantitative ability. *Intelligence, 3*(4), 345-354.

Burnett, S. A., Lane, D. M., & Dratt, L. M. (1982). Spatial ability and handedness. *Intelligence, 6*(1), 57-68.

Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist*, *50*(3), 164-168.

Butovskaya, M. L., & Kozintsev, A. G. (1996). Gender-related factors affecting primate social behavior: Grooming, rank, age, and kinship in heterosexual and all-male groups of stumptail macaques. *American Journal of Physical Anthropology, 101*(1), 39-54.

Campbell, F. A. (1996, June 20-23). *Family factors associated with high and low reading and mathematics scores in children from low income families.* Paper presented at the Head Start National Research Conference, Washington, DC.

Caretta, T. R. (1997). Group differences on U.S. Air Force pilot selection tests. *International Journal of Selection and Assessment, 5*(2), 115-127.

Casey, M. B., & Brabeck, M. M. (1989). Exceptions to the male advantage on a spatial task: Family handedness and college major as factors identifying women who excel. *Neuropsychologia, 27*(5), 689-696.

Casey, M. B., & Brabeck, M. M. (1990). Women who excel on a spatial task: Proposed genetic and environmental factors. *Brain and Cognition, 12*(1), 73-84.

Casey, M. B., Colon, D., & Goris, Y. (1992a). Family handedness as a predictor of mental rotation ability among minority girls in a math-science training program. *Brain and Cognition, 18*(1), 88-96.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, *31*(4), 697-705.

Casey, M. B., & Nuttall, R. L. (1990). Differences in feminine and masculine characteristics in women as a function of handedness: Support for the Geschwind/Galaburda theory of brain organization. *Neuropsychologia, 28*(7), 749-54.

Casey, M. B., Pezaris, E., & Nuttall, R. L. (1992b). Spatial ability as a predictor of math achievement: The imporance of sex and handedness patterns. *Neuropsychologia, 30*(1), 35-45.

Casson, P. R., Elkind-Hirsch, K. E., Buster, J. E., Hornsby, P. J., Carson, S. A., & Snabes, M. C. (1997). Effect of postmenopausal estrogen replacement on circulating androgens. *Obstetrics and Gynecology, 90*(6), 995-998.

Cauley, J. A., Gutai, J. P., Kuller, L. H., LeDonne, D., & Powell, J. G. (1989). The epidemiology of serum sex hormones in postmenopausal women. *American Journal of Epidemiology, 129*(6), 1120-1131.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*(5), 703-722.

Cheung, C. K., & Bagley, C. (1998). Validating an American scale in Hong Kong: The Center for Epidemiological Studies Depression Scale (CES-D). *Journal of Psychology, 132*(2), 169-186.

Cho, M. J., & Kim, K. H. (1998). Use of the Center for Epidemiologic Studies Depression (CES-D) Scale in Korea. *Journal of Nervous and Mental Disease, 186*(5), 304-310.

CHRR. (1995). *NLS handbook: The national longitudinal surveys*. Columbus, OH: Ohio State University.

CHRR. (1996). *NLSY79 codebook supplement: Main file*. Columbus, OH: Center for Human Resource Research.

CHRR. (1997). *NLSY79 users' guide*. Columbus, OH: Ohio State University.

CHRR. (1998). *National longitudinal surveys: Bibliographic data*. Center for Human Resource Research. Available: http://www.chrr.ohio-state.edu/nls-bib.

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants. A proposal. *Archives of General Psychiatry, 44*(6), 573-588.

Cohn, S. J., Cohn, C. M., & Jensen, A. R. (1988). Myopia and intelligence: A pleiotropic relationship? *Human Genetics, 80*(1), 53-58.

Conn, V. S., Taylor, S. G., & Hayes, V. (1992). Social support, self-esteem, and self-care after myocardial infarction. *Health Values, 16*(5), 25-31.

Cooper, C., Kuh, D., Egger, P., Wadsworth, M., & Barker, D. (1996). Childhood growth and age at menarche. *British Journal of Obstetrics and Gynaecology, 103*(8), 814-817.

Copper, R. L., Goldenberg, R. L., DuBard, M. B., & Davis, R. O. (1994). Risk factors for fetal death in white, black, and Hispanic women. Collaborative Group on Preterm Birth Prevention [see comments]. *Obstetrics and Gynecology, 84*(4), 490-495.

Cowchock, F. S., Gibas, Z., & Jackson, L. G. (1993). Chromosome errors as a cause of spontaneous abortion: The relative importance of maternal age and obstetric history. *Fertility and Sterility, 59*(5), 1011-1014.

Crane, J. (1996). Effects of home environment, SES, and maternal test scores on mathematics achievement. *Journal of Educational Research, 89*(5), 305-314.

CRW. (1996). *How schools shortchange girls: The AAUW report*. New York: Marlowe & Company.

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review, 85*(3), 341-364.

Daly, M., & Wilson, M. (1983). *Sex, evolution, and behavior*. Belmont, CA: Wadsworth.

Darwin, C. (1859). *On the origins of species by means of natural selection*. London: John Murray.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

de Ridder, C. M., Bruning, P. F., Zonderland, M. L., Thijssen, J. H., Bonfrer, J. M., Blankenstein, M. A., Huisveld, I. A., & Erich, W. B. (1990). Body fat mass, body fat distribution, and plasma hormones in early puberty in females. *Journal of Clinical Endocrinology and Metabolism, 70*(4), 888-893.

de Ridder, C. M., Thijssen, J. H., Bruning, P. F., Van den Brande, J. L., Zonderland, M. L., & Erich, W. B. (1992). Body fat mass, body fat distribution, and pubertal development: A longitudinal study of physical and hormonal sexual maturation of girls. *Journal of Clinical Endocrinology and Metabolism, 75*(2), 442-446.

Deary, I. A., & Pagliari, C. (1991). The strength of *g* at different levels of ability: Have Detterman and Daniel rediscovered Spearman's 'Law of diminishing returns'? *Intelligence, 15*(2), 247-250.

deBrabander, B., & Boone, C. (1990). Sex differences in perceived locus of control. *Journal of Social Psychology, 130*(2), 271-272.

del Rey, P., & Sheppard, S. (1981). Relationship of psychological androgyny in female athletes to self-esteem. *International Journal of Sport Psychology*, *12*(3), 165-175.

Demo, D. H. (1985). The measurement of self-esteem: Refining our methods. *Journal of Personality and Social Psychology, 48*(6), 1490-1502.

Detterman, D. K. (1991). Reply to Deary and Pagliari: Is *g* intelligence or stupidity? *Intelligence, 15*(2), 251-255.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13*(4), 349-359.

Diamond, J. (1998). *Guns, germs, and steel*. New York: Norton.

Diamond, R., Carey, S., & Back, K. J. (1983). Genetic influences on the development of spatial skills during early adolescence. *Cognition, 13*(2), 167-185.

DMDC. (1998). *ASVAB composites and tests*. Defense Manpower Data Center. [On-line.] Available: http://www.dmdc.osd.mil/asvab.

Doolittle, A. E., & Cleary, T. A. (1987). Gender-based differential item performance in mathematics achievement items. *Journal of Educational Measurement, 24*(2), 157-166.

Dorgan, J. F., Reichman, M. E., Judd, J. T., Brown, C., Longcope, C., Schatzkin, A., Albanes, D., Campbell, W. S., Franz, C., Kahle, L., & et al. (1995). The relation of body size to plasma levels of estrogens and androgens in premenopausal women (Maryland, United States). *Cancer Causes and Control, 6*(1), 3-8.

Dukes, R. L., & Martinez, R. (1994). The impact of ethgender on self-esteem among adolescents. *Adolescence, 29*(113), 105-115.

Eals, M., & Silverman, I. (1994). The hunter-gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethology and Sociobiology*, *15*(2), 95-105.

Earles, J. A., & Ree, M. J. (1992). The predictive validity of the ASVAB for the training grades. *Educational and Psychological Measurement, 52*(3), 721-725.

Eastman, N. J., & Jackson, E. (1968). Weight relationships in pregnancy. I. The bearing of maternal weight gain and pre-pregnancy weight on birth weight in full term pregnancies. *Obstetrics and Gynecological Survey, 23*(11), 1003-1025.

Eaton, S. B., Pike, M. C., Short, R. V., Lee, N. C., Trussell, J., Hatcher, R. A., Wood, J. W., Worthman, C. M., Jones, N. G., Konner, M. J., & et al. (1994). Women's reproductive cancers in evolutionary context. *Quarterly Review of Biology, 69*(3), 353-367.

Ehrhardt, A. A., Epstein, R., & Money, J. (1968). Fetal androgens and female gender identity in the early-treated adrenogenital syndrome. *Johns Hopkins Medical Journal, 122*(3), 160-167.

Entwisle, D. R., & Alexander, K. L. (1996). Family type and children's growth in reading and math over the primary grades. *Journal of Marriage and the Family, 58*(2), 341-355.

ETS. (1998). Unpublished data provided to author .

Evans, M. C. (1998). *On the asymmetry of g*. [On-line]. Available: http://www.mgmt.utoronto.ca-/~evans/iq/g6b.htm.

Eysenck, H. (1981). *A model for personality*. New York: Springer-Verlag.

Fairbank, B. A., Welsh, J. R., & Sawin, L. (1990). *Validity of ASVAB Form 14 for the prediction of high school grades* . San Antonio: Operations Technologies Corporation.

Fava, G. A. (1983). Assessing depressive symptoms across cultures: Italian validation of the CES-D Self-Rating Scale. *Journal of Clinical Psychology, 39*(2), 249-251.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*(2), 95-103.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*(3), 429-456.

Feldman, G. B. (1992). Prospective risk of stillbirth. *Obstetrics and Gynecology, 79*(4), 547-553.

Felson, R. B., & Trudeau, L. (1991). Gender differences in mathematics performance. *Social Psychology Quarterly, 54*(2), 113-126.

Finn, J. (1980). Sex differences in educational outcomes: A cross-national study. *Sex Roles, 6*(1), 9-26.

FNB. (1990). *Recommended dietary allowances* (10th ed.). Washington D.C.: National Research Council.

Forest, M. G. (1983). Role of androgens in fetal and pubertal development. *Hormone Research, 18*(1-3), 69-83.

Frisch, R. E. (1987). Body fat, menarche, fitness and fertility. *Human Reproduction, 2*(6), 521-533.

Frisch, R. E. (1990). The right weight: Body fat, menarche and ovulation. *Baillieres Clinical Obstetrics and Gynaecology, 4*(3), 419-439.

Frisch, R. E., Gotz-Welbergen, A. V., McArthur, J. W., Albright, T., Witschi, J., Bullen, B., Birnholz, J., Reed, R. B., & Hermann, H. (1981). Delayed menarche and amenorrhea of college athletes in relation to age of onset of training. *JAMA, 246*(14), 1559-1563.

Gardner, H., Kornhaber, M. L., & Wake, W. K. (1996). *Intelligence: Multiple perspectives*. New York: Harcourt Brace.

Geary, D. C. (1995). Sexual selection and sex differences in spatial cognition. *Learning and Individual Differences*, *7*(4), 289-301.

Geary, D. C. (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences, 19*(2), 229-246.

Geary, D. C. (1998a). *Male, female: The evolution of human sex differences*. Washington, D.C.: American Psychological Association.

Geary, D. C. (1998b). Sexual selection, the division of labor, and the evolution of sex differences. *Behavioral and Brain Sciences*, *21*(3), 444-448.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67*(5), 2022-2044.

Geary, D. C., & Gilger, J. W. (1989). Age of sexual maturation and adult spatial ability. *Bulletin of the Psychonomic Society, 27*(3), 241-244.

Geary, D. C., Gilger, J. W. & Elliott-Miller, B. (1992). Gender differences in three-dimensional mental rotation: A replication. *Journal of Genetic Psychology 153*(1), 115-117.

Georgiadis, E., Mantzoros, C. S., Evagelopoulou, C., & Spentzos, D. (1997). Adult height and menarcheal age of young women in Greece. *Annals of Human Biology, 24*(1), 55-59.

Geschwind, N., & Behan, P. (1982). Left-handedness: Association with immune disease, migraine, and developmental learning disorder. *Proceedings of the National Academy of Sciences of the United States of America, 79*(16), 5097-5100.

Gilbert, A. N., & Wysocki, C. J. (1992). Hand preference and age in the United States. *Neuropsychologia, 30*(7), 601-608.

Gillis, J. J., DeFries, J. C., & Fulker, D. W. (1992). Confirmatory factor analysis of reading and mathematics performance: A twin study. *Acta Geneticae Medicae et Gemellologiae, 41*(4), 287-300.

Goldenberg, R. L., Hickey, C. A., Cliver, S. P., Gotlieb, S., Woolley, T. W., & Hoffman, H. J. (1997). Abbreviated scale for the assessment of psychosocial status in pregnancy: Development and evaluation. *Acta Obstetricia et Gynecologia Scandinavica Supplementum, 165*, 19-29.

Gouchie, C., & Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology, 16*(4), 323-334.

Grimshaw, G. M., Sitarenios, G., & Finegan, J. A. (1995). Mental rotation at 7 years: Relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition, 29*(1), 85-100.

Hagborg, W. J. (1996). Scores of middle-school-age students on the Rosenberg self-esteem scale. *Psychological Reports, 78*(3, Pt 2), 1071-1074.

Halbreich, U., Lemus, C. Z., Lieberman, J. A., Parry, B., & Schiavi, R. C. (1990). Gonadal hormones, sex and behavior. *Psychopharmacology Bulletin, 26*(3), 297-301.

Halperin, J. M., Newcorn, J. H., Matier, K., Bedi, G., Hall, S., & Sharma, V. (1995). Impulsivity and the initiation of fights in children with disruptive behavior disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*(7), 1199-1211.

Halpern, D. F. (1992). *Sex differences in cognitive abilities*. Hillsdale, NJ: Erlbaum.

Halpern, D. F. (1997). Sex differences in intelligence. Implications for education. *American Psychologist, 52*(10), 1091-1102.

Hamilton, L. S., Nussbaum, E. M., Kupermintz, H., & Kerkhoven, J. I. M. (1995). Enhancing the validity and usefulness of large-scale educational assessments: II. NELS:88 science achievement. *American Educational Research Journal, 32*(3), 555-581.

Hammer, A. L., & Mitchell, W. D. (1996). The distribution of MBTI types in the U.S. by gender and ethnic group. *Journal of Psychological Type, 37*(1), 2-15.

Hampson, E. (1990a). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology, 15*(2), 97-111.

Hampson, E. (1990b). Variations in sex-related cognitive abilities across the menstrual cycle. *Brain and Cognition, 14*(1), 26-43.

Hampson, E. (1995). Spatial cognition in humans: Possible modulation by androgens and estrogens. *Journal of Psychiatry and Neuroscience, 20*(5), 397-404.

Hankinson, S. E., Willett, W. C., Manson, J. E., Hunter, D. J., Colditz, G. A., Stampfer, M. J., Longcope, C., & Speizer, F. E. (1995). Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. *Journal of the National Cancer Institute, 87*(17), 1297-1302.

Hardyck, C., Petrinovich, L. F., & Goldman, R. D. (1976). Left-handedness and cognitive deficit. *Cortex, 12*(3), 266-279.

Harris, J. A., Rushton, J. P., Hampson, E., & Jackson, D. N. (1996). Salivary testosterone and self-report aggressive and pro-social personality characteristics in men and women. *Aggressive Behavior*, *22*(5), 321-331.

Hasegawa, I., Takakuwa, K., & Tanaka, K. (1996). The roles of oligomenorrhoea and fetal chromosomal abnormalities in spontaneous abortions. *Human Reproduction, 11*(10), 2304-2305.

Hassler, M. (1991). Maturation rate and spatial, verbal, and musical abilities: A seven-year longitudinal study. *International Journal of Neuroscience 58*(3-4), 183-198.

Hassler, M., & Nieschlag, E. (1989). Masculinity, feminity, and musical composition: Psychological and psychoendocrinological aspects of musical and spatial faculties. *Archives of Psychology, 141*(1), 71-84.

Hassold, T., Chen, N., Funkhouser, J., Jooss, T., Manuel, B., Matsuura, J., Matsuyama, A., Wilson, C., Yamane, J. A., & Jacobs, P. A. (1980). A cytogenetic study of 1000 spontaneous abortions. *Annals of Human Genetics, 44*(Pt 2), 151-178.

Heath, L., Bresolin, L. B., & Rinaldi, R. C. (1989). Effects of media violence on children. A review of the literature. *Archives of General Psychiatry, 46*(4), 376-379.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41-45.

Held, B. L., Nader, S., Rodriguez-Rigau, L. J., Smith, K. D., & Steinberger, E. (1984). Acne and hyperandrogenism. *Journal of the American Academy of Dermatology, 10*(2 Pt 1), 223-226.

Helm, P., & Lidegaard, O. (1990). The relationship between menarche and sexual, contraceptive, and reproductive life events. *Scandinavian Journal of Primary Health Care, 8*(1), 59-63.

Helm, P., Munster, K., & Schmidt, L. (1995). Recalled menarche in relation to infertility and adult weight and height. *Acta Obstetricia et Gynecologia Scandinavica, 74*(9), 718-722.

Herrmann, W. M., & Beach, R. C. (1978). The psychotropic properties of estrogens. *Pharmakopsychiatrie Neuropsychopharmakologie, 11*(4), 164-176.

Hiernaux, J., & Boedhi Hartono, D. (1980). Physical measurements of the adult Hadza of Tanzania. *Annals of Human Biology, 7*(4), 339-346.

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Petersen (Eds.), *Girls at puberty: Biological and psychosocial perspectives*. New York: Plenum.

Hoffman, L. (1982). Empirical findings concerning sexism in our schools. *Corrective and Social Psychiatry, 28*(3), 100-108.

Holden, C. (1990). Head Start enters adulthood. *Science, 247*(12), 1402.

Holland, J. L. (1972). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice-Hall.

Horn, J. M., & Turner, R. G. (1974). Personality correlates of differential abilities in a sample of lower than average ability. *Psychological Reports, 35*(3), 1211-1220.

Hoyenga, K. B. (1993). Sex differences in human stratification: A biosocial approach. In L. Ellis (Ed.), *Social stratification and socioeconomic inequality: A comparative biosocial analysis* (Vol. 1). Westport, CT: Praeger.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology, 78*(2), 250-261.

Hurtig, A. L., Radhakrishnan, J., Reyes, H. M., & Rosenthal, I. M. (1983). Psychological evaluation of treated females with virilizing congenital adrenal hyperplasia. *Journal of Pediatric Surgery, 18*(6), 887-893.

Jackson, S. A., & Marsh, H. W. (1986). Athletic or antisocial? The female sport experience. *Journal of Sport Psychology*, *83*(3), 198-211.

Jeffe, D. B. (1995). About girls' "difficulties" in science: A social, not a personal matter. *Teachers College Record, 97*(2), 206-226.

Jenkins, J. M., Brook, P. F., Sargeant, S., & Cooke, I. D. (1995). Endocervical mucus pH is inversely related to serum androgen levels and waist to hip ratio. *Fertility and Sterility, 63*(5), 1005-1008.

Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Prager.

Johnson, E. S., & Meade, A. C. (1987). Developmental patterns of spatial ability: An early sex difference. *Child Development*, *58*(3), 725-740.

Jones, J. C., & Barlow, D. H. (1990). Self-reported frequency of sexual urges, fantasies and masturbatory fantasies in heterosexual males and females. *Archives of Sexual Behavior, 19*(3), 269-279.

Jones, M. G. (1989). Gender bias in classroom interactions. *Contemporary Education, 60*(4), 218-222.

Juni, S., Rahamim, E. L., & Brannon, R. (1985). Sex role development as a function of parent models and oedipal fixation. *Journal of Genetic Psychology*, *146*(1), 89-99.

Kaplan, H. S., & Owett, T. (1993). The female androgen deficiency syndrome. *Journal of Sex and Marital Therapy, 19*(1), 3-24.

Karlsson, J. L. (1976). Genetic factors in myopia. *Acta Geneticae Medicae Gemellologiae, 25*, 292-294.

Kass, R. A., Mitchell, K. J., Grafton, F. C., & Wing, H. (1982). *Factor structure of the Armed Services Vocational Aptitude Battery (ASVAB), Forms 8, 9, 10: 1981 Army applicant sample.* Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kaufman, A. S., & McLean, J. E. (1998). An investigation into the relationship between interests and intelligence. *Journal of Clinical Psychology, 54*(2), 279-295.

Kaye, S. A., Folsom, A. R., Soler, J. T., Prineas, R. J., & Potter, J. D. (1991). Associations of body mass and fat distribution with sex hormone concentrations in postmenopausal women. *International Journal of Epidemiology, 20*(1), 151-156.

Kelly, A. (1978). *Girls and science: International study of sex differences in school science achievement*. Stockholm: Almqvist & Wiksell.

Kempner, T. D. (1990). *Social structure and testosterone: Explorations of the socio-bio-social chain.* New Brunswick, NJ: Rutgers.

Kerns, K. A., & Berenbaum, S. A. (1991). Sex differences in spatial ability in children. *Behavioral Genetics, 21*(4), 383-396.

Kimura, D., & Hampson, E. (1993). Neural and hormonal mechanisms mediating sex differences in cognition. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 375-397). Norwood, NJ: Ablex.

Kirchengast, S. (1998). Interactions between somatometric parameters and endogenous hormone levels as well as hormonal induced events in females. *Anthropologischer Anzeiger, 56*(3), 251-265.

Kirchengast, S., Gruber, D., Sator, M., & Huber, J. (1998). Impact of the age at menarche on adult body composition in healthy pre- and postmenopausal women. *American Journal of Physical Anthropology, 105*(1), 9-20.

Kline, P. (1991). *Intelligence: The psychometric view*. London: Routledge.

Knapp, V. J. (1998). Life expectancy, infant mortality and malnutrition in preindustrial Europe: A contemporary explanation. *Nutrition and Health, 12*(2), 89-95.

Knudsen, U. B., Hansen, V., Juul, S., & Secher, N. J. (1991). Prognosis of a new pregnancy following previous spontaneous abortions. *European Journal of Obstetrics, Gynecology, and Reproductive Biology , 39*(1), 31-36.

Kramer, T. R., Moore, R. J., Shippee, R. L., Friedl, K. E., Martinez-Lopez, L., Chan, M. M., & Askew, E. W. (1997). Effects of food restriction in military training on T-lymphocyte responses. *International Journal of Sports Medicine, 18* (Supplement), S84-S90.

Kreiner, D., Liu, H. C., Itskovitz, J., Veeck, L., & Rosenwaks, Z. (1987). Follicular fluid estradiol and progesterone are markers of preovulatory oocyte quality. *Fertility and Sterility, 48*(6), 991-994.

Kunhikrishnan, K., & Manikandan, K. (1992). Sex differences in locus of control: An analysis based on Calicut L.O.C. scale. *Psychological Studies, 37*(2-3), 121-125.

Kunz, J., & Keller, P. J. (1976). HCG, HPL, oestradiol, progesterone and AFP in serum in patients with threatened abortion. *British Journal of Obstetrics and Gynaecology, 83*(8), 640-4.

Kurian, G., & Santhakumari, K. (1994). The effect of sex and familial sinistrality on handedness. *Psychologia: An International Journal of Psychology in the Orient, 37*(1), 30-33.

Lamke, L. K. (1982). The impact of sex-role orientation on self-esteem in early adolescence. *Child Development, 53*(6), 1530-1535.

Lange, A. L., & Fischbein, S. (1990). Qualitative and quantitative aspects of mathematics achievement in MZ and DZ twins. *Acta Geneticae Medicae Gemellologiae, 39*(2), 221-230.

Larsen, S. B., & Kragstrup, J. (1997). [Age at menarche and first sexual intercourse. A cross-sectional study of 17-year old Danish women]. *Ugeskrift for Laeger, 159*(45), 6676-6679.

Le Marchand, L., Kolonel, L. N., Earle, M. E., & Mi, M. P. (1988). Body size at different periods of life and breast cancer risk. *American Journal of Epidemiology, 128*(1), 137-152.

le Roux, A. (1979). Sex differences and the incidence of left-handedness. *Journal of Psychology, 102*(2), 261-262.

Legree, P. J., Pifer, M. E., & Grafton, F. C. (1996). Correlations among cognitive abilities are lower for higher ability groups. *Intelligence, 23*(1), 45-57.

Lehrke, R. (1997). *Sex linkage of intelligence: The x factor*. Westport, CT: Praeger.

Leondari, A., Syngollitou, E., & Kiosseoglou, G. (1998). Academic achievement, motivation and future selves. *Educational Studies, 24*(2), 153-163.

Levin, J. S., Taylor, R. J., & Chatters, L. M. (1994). Race and gender differences in religiosity among older adults: Findings from four national surveys. *Journal of Gerontology, 49*(3), S137-S145.

Levy, P. E., & Baumgardner, A. H. (1991). Effects of self-esteem and gender on goal choice. *Journal of Organizational Behavior, 12*(6), 529-541.

Lin, L. L., & Chen, C. J. (1987). Twin study on myopia. *Acta Geneticae Medicae et Gemellologiae, 36*(4), 535-540.

Lippa, R. (1998a). Gender-related individual differences and National Merit Test performance: Girls who are "masculine" and boys who are "feminine" tend to do better. In L. Ellis & L. Ebertz (Eds.), *Males, females, and behavior: Toward biological understanding.* (pp. 177-193). Westport, CT: Praeger.

Lippa, R. (1998b). Gender-related individual differences and the structure of vocational interests: The importance of the people-things dimension. *Journal of Personality and Social Psychology*, *74*(4), 996-1009.

Locurto, C. (1990). The malleability of IQ as judged from adoption studies. *Intelligence, 14*, 275-292.

Loesch, D. Z., Huggins, R., Rogucka, E., Hoang, N. H., & Hopper, J. L. (1995). Genetic correlates of menarcheal age: A multivariate twin study. *Annals of Human Biology , 22*(6), 470-490.

Longcope, C. (1986). Adrenal and gonadal androgen secretion in normal females. *Clinics in Endocrinology and Metabolism, 15*(2), 213-228.

Lopez, L. C., & Staszkiewicz, M. J. (1985). Sex differences in internality-externality. *Psychological Reports, 57*(3), 1159-1164.

Loret De Mola, J. R., Garside, W. T., Bucci, J., Tureck, R. W., & Heyner, S. (1997). Analysis of the human zona pellucida during culture: Correlation with diagnosis and the preovulatory hormonal environment. *Journal of Assisted Reproduction and Genetics, 14*(6), 332-336.

Lorr, M., & Wunderlich, R. A. (1986). Two objective measures of self-esteem. *Journal of Personality Assessment, 50*(1), 18-23.

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math/science pipeline. *Current Directions in Psychological Science, 1*(2), 61-66.

Lubinski, D., & Humphreys, L. G. (1990). A broadly based analysis of mathematical giftedness. *Intelligence, 14*(3), 327-355.

Lubinski, D., & Humphreys, L. G. (1992). Some bodily and medical correlates of mathematical giftedness and commensurate levels of socioeconomic status. *Intelligence, 16*(1), 99-115.

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26*(2), 254-263.

Lynn, R. (1987). The intelligence of the Mongoloids: A psychometric, evolutionary, and neurological theory. *Personality and Individual Differences, 8*(6), 813-844.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin, 109*(2), 267-296.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

MacDonald, K. (1994). *A people that shall dwell alone: Judaism as a group evolutionary strategy*. Westport, CT: Praeger.

Madigan, M. P., Troisi, R., Potischman, N., Dorgan, J. F., Brinton, L. A., & Hoover, R. N. (1998). Serum hormone levels in relation to reproductive and lifestyle factors in postmenopausal women (United States). *Cancer Causes and Control, 9*(2), 199-207.

Malina, R. M. (1996). Tracking of physical activity and physical fitness across the lifespan. *Research Quarterly for Exercise and Sport, 67*(3 Supplement), S48-S57.

Marshall, S. P., & Smith, J. D. (1987). Sex differences in learning mathematics: A longitudinal study with item and error analyses. *Journal of Educational Psychology, 79*(4), 372-383.

Martino, G., & Winner, E. (1995). Talents and disorders: Relationships among handedness, sex, and college major. *Brain and Cognition, 29*(1), 66-84.

Mathur, R. S., Landgrebe, S. C., Moody, L. O., Semmens, J. P., & Williamson, H. O. (1985). The effect of estrogen treatment on plasma concentrations of steroid hormones, gonadotropins, prolactin and sex hormone-binding globulin in post-menopausal women. *Maturitas, 7*(2), 129-133.

Matthews, M. H. (1987). Sex differences in spatial competence: The ability of young children to map "primed" unfamiliar environments. *Educational Psychology*, *7*(2), 77-90.

McCallum, D. I., Smith, I. M., & Eliot, J. (1979). Further investigation of components of mathematical ability. *Psychological Reports, 44*(3, Pt 2), 1127-1133.

McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., & et al. (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science, 276*(5318), 1560-1563.

McFarland, R. (1997). Female primates: Fit or fat? In M. E. Morbeck, A. Galloway, & A. L. Zihlman (Eds.), *The evolving female: A life history perspective* (pp. 163-175). Princeton: Princeton University Press.

McQueen, J. K., Wilson, H., & Fink, G. (1997). Estradiol-17 beta increases serotonin transporter (SERT) mRNA levels and the density of SERT-binding sites in female rat brain. *Brain Research. Molecular Brain Research, 45*(1), 13-23.

Mead, L. A., & Hampson, E. (1996). A sex difference in turning bias in humans. *Behavioural Brain Research, 78*(2), 73-79.

Measurement-Group. (1997). *Current psychological distress form.* [On-line]. Available: http://www.tmg-web.com/modules/module26.-htm.

Meece, J., & Jones, G. (1996). Girls in mathematics and science: Constructivism as a feminist perspective. *High School Journal, 79*(3), 242-248.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Student/teacher relations and attitudes toward mathematics before and after the transition to junior high school. *Child Development, 60*(4), 981-992.

Miller, A. S., & Hoffman, J. P. (1995). Risk and religion: An explanation of gender differences in religiosity. *Journal for the Scientific Study of Religion, 34*(1), 63-75.

Miller, E. M. (1992). On the correlation of myopia and intelligence. *Genetic, Social, and General Psychology Monographs, 118*(4), 361-383.

Mills, C. J. (1998). Personality, learning style, and cognitive style profiles of mathematically talented students. *European Journal for High Ability, 9*(1), 70-85.

Mills, C. J., Ablard, K. E., & Stumpf, H. (1993). Gender differences in academically talented young students' mathematical reasoning: Patterns across age and subskills. *Journal of Educational Psychology*, *85*(2), 340-346.

Mittendorf, R., Longnecker, M. P., Newcomb, P. A., Dietz, A. T., Greenberg, E. R., Bogdan, G. F., Clapp, R. W., & Willett, W. C. (1995). Strenuous physical activity in young adulthood and risk of breast cancer (United States). *Cancer Causes and Control, 6*(4), 347-53.

Moffat, S. D., & Hampson, E. (1996). A curvilinear relationship between testosterone and spatial cognition in humans: Possible influence of hand preference. *Psychoneuroendocrinology, 21*(3), 323-337.

Moore, C. L. (1986). A hormonal basis for sex differences in the self-grooming of rats. *Hormones and Behavior, 20*(2), 155-165.

Moore, S. D., & Stanley, J. C. (1987). Family backgrounds of young Asian Americans who reason extremely well mathematically. Baltimore: Johns Hopkins University.

Morris, N. M., Udry, J. R., Khan-Dawood, F., & Dawood, M. Y. (1987). Marital sex frequency and midcycle female testosterone. *Archives of Sexual Behavior, 16*(1), 27-37.

Myers, A. M., & Lips, H. M. (1978). Participation in competitive amateur sports as a function of psychological androgyny. *Sex Roles*, *4*(4), 571-578.

NAEP. (1996). *Report in brief: NAEP 1996 trends in academic progress*. [On-line]. Available: http://nces.ed.gov/naep/96report/-97986f4s.html.

Nagata, C., Kaneda, N., Kabuto, M., & Shimizu, H. (1997). Factors associated with serum levels of estradiol and sex hormone- binding globulin among premenopausal Japanese women. *Environmental Health Perspectives, 105*(9), 994-997.

NLSY79-- National Longitudinal Survey of Youth 1979-1996 (Release 9.0) [CD-ROM]. (1996). Columbus, OH: Center for Human Resource Research [Producer and Distributor].

Natsopoulos, D., Kiosseoglou, G., & Xeromeritou, A. (1992). Handedness and spatial ability in children: Further support for Geschwind's hypothesis of "pathology of superiority" and for Annett's theory of intelligence. *Genetic, Social, and General Psychology Monographs, 118*(1), 103-126.

NSF. (1994). Women, minorities, and persons with disabilities in science and engineering. Washington, D.C.: National Science Foundation.

Nunn, G. D. (1994). Adult learners' locus of control, self-evaluation, and learning temperament as a function of age and gender. *Journal of Instructional Psychology, 21*(3), 260-264.

Nyborg, H. (1983). Spatial ability in men and women: Review and new theory. *Advances in Behavior Research and Therapy, 5*(2), 89-140.

Nyborg, H. (1984). Performance and intelligence in hormonally different groups. In G. J. Devries, J. P. C. Bruin, H. B. M. Uylings, & M. A. Corner (Eds.), *Sex differences in the brain: Relation between structure and function* (pp. 491-508). Amsterdam: Elsevier.

Nyborg, H. (1988). Mathematics, sex hormones, and brain function. *Behavioral and Brain Sciences, 11*(2), 206-207.

Nyborg, H. (1994). *Hormones, sex, and society*. Westport, CT: Praeger.

Ogunyemi, D., Jackson, U., Buyske, S., & Risk, A. (1998). Clinical and pathologic correlates of stillbirths in a single institution. *Acta Obstetricia Gynecologia Scandinavica, 77*(7), 722-728.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Orme, J. G., Reis, J., & Herz, E. J. (1986). Factorial and discriminant validity of the Center for Epidemiological Studies Depression (CES-D) scale. *Journal of Clinical Psychology, 42*(1), 28-33.

Ortiz, V., & Cooney, R. S. (1984). Sex-role attitudes and labor force participation among young Hispanic females and non-Hispanic white females. *Social Science Quarterly, 65*(2), 392-400.

Overholser, J. C., Schubert, D. S., Foliart, R., & Frost, F. (1993). Assessment of emotional distress following a spinal cord injury. *Rehabilitation Psychology, 38*(3), 187-198.

Pamich, A. (1987). Personalita della donna atleta. *Movimento*, *3*(1), 63-65.

Parazzini, F., Chatenoud, L., Tozzi, L., Benzi, G., Dal Pino, D., & Fedele, L. (1997). Determinants of risk of spontaneous abortions in the first trimester of pregnancy. *Epidemiology, 8*(6), 681-683.

Parazzini, F., la Vecchia, C., Negri, E., Bruzzi, P., Palli, D., & Boyle, P. (1990). Anthropometric variables and risk of breast cancer. *International Journal of Cancer, 45*(3), 397-402.

Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development, 53*(2), 310-321.

Pearlin, L. I., Aneshensel, C. S., & LeBlanc, A. J. (1997). The forms and mechanisms of stress proliferation: The case of AIDS caregivers. *Journal of Health and Social Behavior, 38*(3), 223-236.

Pearlin, L. I., Lieberman, M. A., Menaghan, E. G., & Mullen, J. T. (1981). The stress process. *Journal of Health and Social Behavior, 22*, 337-356.

Pearlin, L. I., & Skaff, M. M. (1996). Stress and the life course: A paradigmatic alliance. *Gerontologist, 36*(2), 239-247.

Persky, H., Lief, H. I., Strauss, D., Miller, W. R., & O'Brien, C. P. (1978). Plasma testosterone level and sexual behavior of couples. *Archives of Sexual Behavior, 7*(3), 157-173.

Peters, M. (1991). Sex, handedness, mathematical ability, and biological causation. *Canadian Journal of Psychology, 45*(3), 415-419.

Petersen, A. C. (1976). Physical androgyny and cognitive functioning in adolescence. *Developmental Psychology*, *12*(6), 524-533.

Phelps, J. Y., Levine, A. S., Hickman, T. N., Zacur, H. A., Wallach, E. E., & Hinton, E. L. (1998). Day 4 estradiol levels predict pregnancy success in women undergoing controlled ovarian hyperstimulation for IVF. *Fertility and Sterility, 69*(6), 1015-1019.

Phillips, K., & Silverman, I. (1997). Differences in the relationship of menstrual cycle phase to spatial performance on two- and three-dimensional tasks. *Hormones and Behavior, 32*(3), 167-175.

Pi-Sunyer, X. (1988). Obesity. In J. B. Wyngaarden & L. H. Smith (Eds.), *Cecil textbook of medicine* (18th ed., pp. 1219-1228). Philadelphia: W.B. Saunders.

Plomin, R., & Thompson, L. A. (1993). Genetics and high cognitive ability. *Ciba Foundation Symposium, 178*, 67-79.

Plotnick, R. D., & Butler, S. S. (1991). Attitudes and adolescent nonmarital childbearing: Evidence from the National Longitudinal Survey of Youth. *Journal of Adolescent Research, 6*(4), 470-492.

Purifoy, F. E. (1981). Endocrine-environment interaction in human variability. *Annual Review of Anthropology*, *10*(1), 141-162.

Radloff, L. S. (1977). The CES-D scale. A self report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Rantala, M. L., Stenman, U. H., & Koskimies, A. I. (1988). Serum androgens in infertile women with ovulatory dysfunction. *Human Reproduction, 3*(4), 437-439.

Ratliff, K. G., & Burkhart, B. R. (1984). Sex differences in motivations for and effects of drinking among college students. *Journal of Studies on Alcohol, 45*(1), 26-32.

Ray, W. J., Newcombe, N., Semon, S., & Cole, P. M. (1981). Spatial abilities, sex differences, and EEG functioning. *Neuropsychologia, 19*(5), 719-722.

Raymond, C. L., & Benbow, C. P. (1986). Gender differences in mathematics: A function of parental support and student sex typing? *Developmental Psychology*, *22*(6), 808-819.

Rebar, R. W. (1988). The ovaries. In J. B. Wyngaarden & L. H. Smith (Eds.), *Cecil textbook of medicine* (18th ed., pp. 1425-1446). Philadelphia: W.B. Saunders.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of the ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement, 54*(2), 459-463.

Reiss, M., Reiss, G., & Freye, H. A. (1998). Some aspects of self-reported hand preference. *Perceptual and Motor Skills, 86*(3 Pt 1), 953-954.

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology*, *22*(2), 191-198.

Ricalde, A. E., Velasquez-Melendez, G., Tanaka, A. C., & de Siqueira, A. A. (1998). Mid-upper arm circumference in pregnant women and its relation to birth weight. *Revista de Saude Publica, 32*(2), 112-117.

Richman, C. L., Clark, M. L., & Brown, K. P. (1985). General and specific self-esteem in late adolescent students: Race, gender, SES effects. *Adolescence, 20*(79), 555-566.

Rierdan, J., & Koff, E. (1984). Age at menarche and cognitive functioning. *Bulletin of the Psychonomic Society, 22*(3), 174-176.

Roberts, R. E., Vernon, S. W., & Rhoades, H. M. (1989). Effects of language and ethnic status on reliability and validity of the Center for Epidemiologic Studies-Depression Scale with psychiatric patients. *Journal of Nervous and Mental Disease, 177*(10), 581-592.

Rockhill, B., Moorman, P. G., & Newman, B. (1998). Age at menarche, time to regular cycling, and breast cancer. *Cancer Causes and Control, 9*(4), 447-453.

Rodriguez, M. M., & Grossberg, G. T. (1998). Estrogen as a psychotherapeutic agent. *Clinics in Geriatric Medicine, 14*(1), 177-189.

Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, NJ: Princeton University Press.

Rovet, J., Szekely, C., & Hockenberry, M. N. (1994). Specific arithmetic calculation deficits in children with Turner syndrome. *Journal of Clinical and Experimental Neuropsychology, 16*(6), 820-839.

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior.* New York: Guilford Press.

Sadker, M., & Sadker, D. (1986). Sexism in the classroom: From grade school to graduate school. *Phi Delta Kappan, 67*(7), 512-515.

Sanders, B., & Soares, M. P. (1986). Sexual maturation and spatial ability in college students. *Developmental Psychology, 22*(2), 199-203.

Sandler, S. W. (1977). Spontaneous abortion in perspective. A 7-year study. *South African Medical Journal, 52*(28), 1115-1118.

Sandqvist, K. (1995). Verbal boys and mathematical girls: Family background and educational careers. *Scandinavian Journal of Educational Research, 39*(1), 5-36.

Schulberg, H. C., Saul, M., McClelland, M., Ganguli, M., Christy, W., & Frank, R. (1985). Assessing depression in primary medical and psychiatric practices. *Archives of General Psychiatry, 42*(12), 1164-1170.

Seidman, S. N., & Rieder, R. O. (1994). A review of sexual behavior in the United States. *American Journal of Psychiatry, 151*(3), 330-341.

Selkow, P. (1987). Male/female differences in mathematical ability: A function of biological sex or perceived gender role? *Psychological Reports, 57*(2), 551-557.

Shangold, M. M., Kelly, M., Berkeley, A. S., Freedman, K. S., & Groshen, S. (1989). Relationship between menarcheal age and adult height. *Southern Medical Journal, 82*(4), 443-445.

Shapiro, D., & Crowley, J. E. (1981). Aspirations and expectations for the future: Education, work activity, and fertility. In M. E. Borus (Ed.), *Pathways to the future: A report on the National Longitudinal Survey of Youth Labor Market Experience in 1979.* . Columbus: Ohio State University.

Shepard, P. (1998). *Coming home to the Pleistocene*. Washington D.C.: Island Press.

Sherwin, B. B., & Gelfand, M. M. (1987). The role of androgen in the maintenance of sexual functioning in oophorectomized women. *Psychosomatic Medicine, 49*(4), 397-409.

Signorella, M. L., & Jamison, W. (1986). Masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin, 100*(2), 207-228.

Silverman, I., & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* . New York: Oxford.

Silverman, I., & Phillips, K. (1993). Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethology and Sociobiology*, *14*(4), 257-269.

Silverman, I., & Phillips, K. (1998). Psychology of spatial sex differences. In C. Crawford & D. L. Krebs (Eds.), *Handbook of evolutionary psychology: Ideas, issues, and applications* (pp. 595-612). Mahwah NJ: Erlbaum.

Silverman, I., Phillips, K., & Silverman, L. K. (1996). Homogeneity of effect sizes for sex across spatial tests and cultures: Implications for hormonal theories. *Brain and Cognition, 31*(1), 90-94.

Sizonenko, P. C. (1989). Physiology of puberty. *Journal of Endocrinological Investigation, 12*(8), 59-63.

Skaff, M. M., & Pearlin, L. I. (1992). Caregiving: Role engulfment and the loss of self. *Gerontologist, 32*(5), 656-664.

Skaff, M. M., Pearlin, L. I., & Mullan, J. T. (1996). Transitions in the caregiving career: Effects on sense of mastery. *Psychology and Aging, 11*(2), 247-257.

Smith, B. D., Meyers, M. B., & Kline, R. (1989). For better or for worse: Left-handedness, pathology, and talent. *Journal of Clinical and Experimental Neuropsychology, 11*(6), 944-958.

Smith, K. D., Rodriguez-Rigau, L. J., Tcholakian, R. K., & Steinberger, E. (1979). The relation between plasma testosterone levels and the lengths of phases of the menstrual cycle. *Fertility and Sterility, 32*(4), 403-407.

Sommers, C. H. (1994). *Who stole feminism: How women have betrayed women*. New York: Touchstone.

Stanley, J. (1993). Boys and girls who reason well mathematically. *Ciba Foundation Symposium, 178*, 119-138.

Starr, B. S. (1979). Sex differences among personality correlates of mathematical ability in high school seniors. *Psychology of Women Quarterly, 4*(2), 212-220.

Steedman, J. (1985). Examination results in mixed and single-sex schools. In D. Reynolds (Ed.), *Studying school effectiveness* . London: Falmer.

Steen, R. G. (1996). *DNA and destiny: Nature and nurture in human behavior*. New York: Plenum.

Steffenhagen, L. A., & Steffenhagen, R. A. (1985). Self-esteem and primary demographic characteristics of alcoholics in a rural state. *Journal of Alcohol and Drug Education, 30*(2), 51-59.

Steinberger, E., Rodriguez-Rigau, L. J., Smith, K. D., & Held, B. (1981). The menstrual cycle and plasma testosterone levels in women with acne. *Journal of the American Academy of Dermatology, 4*(1), 54-58.

Steinberger, E., Smith, K. D., Tcholakian, R. K., & Rodriguez-Rigau, L. J. (1979). Testosterone levels in female partners of infertile couples. Relationship between androgen levels in the woman, the male factor, and the incidence of pregnancy. *American Journal of Obstetrics and Gynecology, 133*(2), 133-138.

Stevenson, H. W., Lee, S. Y., Chen, C. S., Lummis, M., Stigler, J., Fan, L., & Ge, F. (1990). Mathematics achievement of children in China and the United States. *Child Development, 61*(4), 1053-1066.

Storfer, M. D. (1990). *Intelligence and giftedness*. San Francisco: Jossey-Bass.

Stumpf, H. (1995). Gender differences in performance on tests of cognitive abilities: Experimental design issues and empirical results. *Learning and Individual Differences, 7*(4), 275-287.

Stunkard, A. J., Foch, T. T., & Hrubec, Z. (1986). A twin study of human obesity. *JAMA, 256*(1), 51-54.

Svrakic, D. M., Przybeck, T. R., & Cloninger, C. R. (1991). Further contribution to the conceptual validity of the unified biosocial model of personality: U.S. and Yugoslav data. *Comprehensive Psychiatry, 32*(3), 195-209.

Swanson, C. A., Brinton, L. A., Taylor, P. R., Licitra, L. M., Ziegler, R. G., & Schairer, C. (1989). Body size and breast cancer risk assessed in women participating in the Breast Cancer Detection Demonstration Project. *American Journal of Epidemiology, 130*(6), 1133-1141.

Takeuchi, T., Nishii, O., Okamura, T., Yaginuma, T., & Kawana, T. (1993). Free testosterone and abortion in early pregnancy. *International Journal of Gynaecology and Obstetrics, 43*(2), 151-156.

Teikari, J. M., Kaprio, J., Koskenvuo, M. K., & Vannas, A. (1988). Heritability estimate for refractive errors--a population-based sample of adult twins. *Genetic Epidemiology, 5*(3), 171-181.

Telama, R., Leskinen, E., & Yang, X. (1996). Stability of habitual physical activity and sport participation: A longitudinal tracking study. *Scandinavian Journal of Medicine and Science in Sports, 6*(6), 371-378.

Temple, C. M. (1990). Academic discipline, handedness and immune disorders. *Neuropsychologia, 28*(3), 303-308.

Thor, D. H., Harrison, R. J., Schneider, S. R., & Carr, W. J. (1988). Sex differences in investigatory and grooming behaviors of laboratory rats (Rattus norvegicus) following exposure to novelty. *Journal of Comparative Psychology, 102*(2), 188-192.

Tocci, C. M., & Englehard, G. (1991). Achievement, parental support, and gender differences in attitudes toward mathematics. *Journal of Educational Research, 84*(5), 280-286.

Toner, J. P., & Flood, J. T. (1993). Fertility after the age of 40. *Obstetrics and Gynecology Clinics of North America, 20*(2), 261-272.

Tooby, J., & Cosmides, L. (1989). Evolutionary psychology and the generation of culture, part 1. *Ethology and Sociobiology, 10*(1-3), 24-49.

Udry, J. R. (1980). Changes in the frequency of marital intercourse from panel data. *Archives of Sexual Behavior, 9*(4), 319-325.

Udry, J. R., Morris, N. M., & Kovenock, J. (1995). Androgen effects on women's gendered behaviour. *Journal of Biosocial Science, 27*(3), 359-368.

van Lunsen, R. H., & Laan, E. (1997). Sex, hormones and the brain. *European Journal of Contraception and Reproductive Health Care, 2*(4), 247-251.

Vanderzanden, J. W. (1989). *Human development*. New York: Knopf.

VanGoozen, S. H. M. (1994). *Male and female: Effects of sex hormones on aggression, cognition, and sexual motivation*. Amsterdam: University of Amsterdam.

VanGoozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & VanDePol, N. E. (1995). Gender differences in behavior: Activating effects of cross-sex hormones. *Psychneuroendocrinology, 20*(8), 343-363.

Vatten, L. J., & Kvinnsland, S. (1990). Body height and risk of breast cancer. A prospective study of 23,831 Norwegian women. *British Journal of Cancer, 61*(6), 881-885.

Vernon, P. A. (1950). *The structure of human abilities*. London: Methuen.

Vernon, P. A. (1993). *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.

Vernon, P. A. (Ed.). (1994). *The neuropsychology of individual differences*. New York: Academic Press.

Vihko, R., & Apter, D. (1984). Endocrine characteristics of adolescent menstrual cycles: Impact of early menarche. *Journal of Steroid Biochemistry, 20*(1), 231-236.

Waber, D. P. (1977). Sex differences in mental abilities, hemispheric lateralization, and rate of physical growth at adolescence. *Developmental Psychology, 13*(1), 29-38.

Waber, D. P., Mann, M. B., Merola, J., & Moylan, P. M. (1985). Physical maturation rate and cognitive performance in early adolescence: A longitudinal examination. *Developmental Psychology*, *21*(4), 666-681.

Walford-Kraemer, P., & Light, H. K. (1984). Depression and mastery in women: Differences according to personal characteristics. *Psychological Reports, 54*(3), 710.

Warnock, J. K., Bundren, J. C., & Morris, D. W. (1997). Female hypoactive sexual desire disorder due to androgen deficiency: Clinical and psychometric issues. *Psychopharmacology Bulletin, 33*(4), 761-766.

Weigley, E. S. (1984). Average? Ideal? Desirable? A brief overview of height-weight tables in the United States. *Journal of the American Dietetic Association, 84*(4), 417-423.

Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance in adolescence. *Intelligence, 16*(1), 117-135.

Weissman, M. M., Sholomskas, D., Pottenger, M., Prusoff, B. A., & Locke, B. Z. (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology, 106*(3), 203-214.

Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): Integrative review of validity studies*. San Antonio, TX: Operational Technologies Corporation.

Wilder, G. Z. (1996). *Correlates of gender differences in cognitive functioning*. New York: College Entrance Examination Board.

Williams, C. L., & Meck, W. H. (1991). The organizational effects of gonadal steroids on sexually dimorphic spatial ability. *Psychoneuroendocrinology, 16*(1-3), 155-176.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah, NJ: Erlbaum.

Wilson, E. O. (1980). *Sociobiology: The abridged edition*. Cambridge, MA: Belknap.

Wilson, G. (1992). *The great sex divide*. Washington, D.C.: Scott-Townsend.

Wilson, G. D., & Lang, R. J. (1981). Sex differences in sexual fantasy patterns. *Personality and Individual Differences, 2*(4), 343-346.

Wong, N. Y. (1992). The relationship between mathematical achievement, affective variables, and home background. *Mathematics Education Research Journal, 4*(3), 32-42.

Wooten, B. H. (1997, April). Gender differences in occupational employment. *Monthly Labor Review*, pp. 15-24.

Wramsby, H., Kullander, S., Liedholm, P., Rannevik, G., Sundstrom, P., & Thorell, J. (1981). The success rate of in vitro fertilization of human oocytes in relation to the concentrations of different hormones in follicular fluid and peripheral plasma. *Fertility and Sterility, 36*(4), 448-454.

Yoo, K. Y., Kim, H., Shin, H. R., Kang, D., Ha, M., Park, S. K., Lee, B. O., & Cho, S. H. (1998). Female sex hormones and body mass in adolescent and postmenopausal Korean women. *Journal of Korean Medical Science, 13*(3), 241-246.

Zaadstra, B. M., Seidell, J. C., Van Noord, P. A., te Velde, E. R., Habbema, J. D., Vrieswijk, B., & Karbaat, J. (1993). Fat and female fecundity: Prospective study of effect of body fat distribution on conception rates. *BMJ, 306*(6876), 484-487.

Zehr, J. L., Maestripieri, D., & Wallen, K. (1998). Estradiol increases female sexual initiation independent of male responsiveness in rhesus monkeys. *Hormones and Behavior, 33*(2), 95-103.

Zumoff, B. (1988). Hormonal abnormalities in obesity. *Acta Medica Scandinavica Supplementum, 723*, 153-160.