

Gender Differences in Mathematics

An Integrative Psychological Approach

Edited by

ANN M. GALLAGHER

Law School Admission Council

JAMES C. KAUFMAN

California State University, San Bernadino



PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Cambridge University Press 2005

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2005

Printed in the United States of America

Typeface Palatino 10/12 pt. *System* L^AT_EX 2_ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Gender differences in mathematics / edited by Ann Gallagher, James Kaufman.

p. cm.

Includes bibliographical references and index.

ISBN 0-521-82605-5 – ISBN 0-521-53344-9 (pbk.)

1. Mathematical ability – Sex differences. 2. Mathematics – Study and
teaching – Psychological aspects. 3. Sex differences in education.

I. Gallagher, Ann M. II. Kaufman, James C.

QA27.5.G47 2004

510'.8 – dc22

2003069686

ISBN 0 521 82605 5 hardback

ISBN 0 521 53344 9 paperback

Contents

<i>Preface</i>	page ix
<i>Acknowledgments</i>	xiii
<i>List of Contributors</i>	xv
1 Research on the Women and Mathematics Issue: A Personal Case History <i>Susan F. Chipman</i>	01
2 The Perseverative Search for Sex Differences in Mathematics Ability <i>Jeremy B. Caplan and Paula J. Caplan</i>	25
3 A Psychobiosocial Model: Why Females Are Sometimes Greater Than and Sometimes Less Than Males in Math Achievement <i>Diane F. Halpern, Jonathan Wai, and Amanda Saw</i>	48
4 Gender Differences in Math: Cognitive Processes in an Expanded Framework <i>James P. Byrnes</i>	73
5 Cognitive Contributions to Sex Differences in Math Performance <i>James M. Royer and Laura M. Garofoli</i>	99
6 Spatial Ability as a Mediator of Gender Differences on Mathematics Tests: A Biological–Environmental Framework <i>Ronald L. Nuttal, M. Beth Casey, and Elizabeth Pezaris</i>	121
7 Examining Gender-Related Differential Item Functioning Using Insights from Psychometric and Multicontext Theory <i>Allan S. Cohen and Robert A. Ibarra</i>	143
	vii

8	The Gender-Gap Artifact: Women's Underperformance in Quantitative Domains Through the Lens of Stereotype Threat <i>Paul G. Davies and Steven J. Spencer</i>	172
9	"Math is hard!" (Barbie™, 1994): Responses of Threat vs. Challenge-Mediated Arousal to Stereotypes Alleging Intellectual Inferiority <i>Talia Ben-Zeev, Cristina M. Carrasquillo, Alison M. L. Ching, Tattiya J. Kliengklom, Kristen L. McDonald, Daniel C. Newhall, Gillian E. Patton, Tiffany D. Stewart, Tonya Stoddard, Michael Inzlicht, and Steven Fein</i>	189
10	The Role of Ethnicity on the Gender Gap in Mathematics <i>Alyssa M. Walters and Lisa M. Brown</i>	207
11	The Gender Gap in Mathematics: Merely a Step Function? <i>Sophia Catsambis</i>	220
12	"I can, but I don't want to": The Impact of Parents, Interests, and Activities on Gender Differences in Math <i>Janis E. Jacobs, Pamela Davis-Kean, Martha Bleeker, Jacquelynne S. Eccles, and Oksana Malanchuk</i>	246
13	Gender Effects on Mathematics Achievement: Mediating Role of State and Trait Self-Regulation <i>Eunsook Hong, Harold F. O'Neil, and David Feldon</i>	264
14	Gender Differences in Mathematics Self-Efficacy Beliefs <i>Frank Pajares</i>	294
15	Gender Differences in Mathematics: What We Know and What We Need to Know <i>Ann M. Gallagher and James C. Kaufman</i>	316
	<i>Author Index</i>	333
	<i>Subject Index</i>	345

Research on the Women and Mathematics Issue

A Personal Case History

Susan F. Chipman

The history of research on the issue of women's participation in mathematics provides an interesting case study of the psychology and sociology of research in the social sciences. Although there had been prior research on the topic, two key works of the early and mid-1970s sparked a major burst of interest. They were Lucy Sell's unpublished study of women at the University of California at Berkeley (Sells, 1973), "High school mathematics as the critical factor in the job market," and Sheila Tobias's publications on math anxiety (Tobias, 1976, 1978), the first of them an article in *MS* magazine in 1976. The study of mathematics, or the failure to study mathematics, came to be seen as a critical barrier to women's participation in a wide range of high-status and remunerative occupations during those surging years of the women's movement. Based on a random sample of freshmen entering Berkeley in 1972, Sells (1973) reported that only 8% of the females had taken four years of high school mathematics, whereas 57% of the men had. This report received a lot of attention.

The U.S. National Institute of Education (NIE) responded with plans for a special grants competition addressing this perceived problem. Background preparations for this competition were exceptionally thorough. Three review papers were commissioned to examine existing research results and opinions concerning major classes of possible influences on women's choices to study mathematics or to select occupations requiring mathematical competence: Fennema (1977) reviewed cognitive, affective, and educational influences; Fox (1977) reviewed social influences; and Sherman (1977) reviewed possible biological explanations. These papers were presented at a large, 2-day-long working conference in Washington, DC, that brought together many people concerned with the mathematics education of women, in February 1977. A grants announcement was issued (NIE, 1977). The research grants were intended to provide "a better knowledge base for designing effective educational programs to encourage women to enroll in mathematics beyond the minimal school

requirements." An important underlying assumption was expressed in the opening statement describing the research requested by the announcement, "Women's lower enrollment in the study of advanced mathematics precludes them from entering a variety of occupations requiring mathematical competence."

The grants competition was sponsored by an organizational unit called the Career Awareness Division of the Education and Work Group of the NIE. By the time the research projects were completed, there had been a major reorganization of the NIE. I found myself responsible for this research program, and for a planned publication to pull the research results together, because they had been grouped with all other research on mathematics learning, in a division on Learning and Development that I was chosen to direct. The planned summary publication for the research program was to include chapters by each supported researcher as well as a research synthesis. Although my earlier involvement in the grants competition had been somewhat peripheral – I had attended the working conference and had served as a reviewer of grant proposals – I chose to take on the job of synthesizing the research myself, rather than contracting it out, as originally planned (Chipman, Brush, & Wilson, 1985; Chipman & Thomas, 1985; Chipman & Wilson, 1985). At the NIE, we were continuing to receive more grants proposals on the topic of women (or girls) and mathematics than on all other topics in mathematics education combined. This seemed disproportionate. Mathematics education was not, and still is not, a well-researched area. Many problems concerning more effective ways to teach mathematics had not been addressed. It was part of my job responsibility to define and set research priorities.

In this chapter, I discuss how I have come to understand the women and mathematics issue since the late 1970s, in all its many dimensions. I have revisited the issue many times (Chipman, 1994; Chipman, 1996a, 1996b), sometimes also considering related issues such as participation in fields of science and technology and the participation of minorities, with separate consideration of minority women (Chipman & Thomas, 1987). In addition to these review efforts, I have pursued some research into specific aspects of the issue: possible test bias (Chipman, 1988b; Chipman, Marshall, & Scott, 1991) and the impact of mathematics anxiety on choice of major field and career (Chipman, Krantz, & Silver, 1992, 1995).

As I began the task of synthesizing the set of research grants on women and mathematics, it seemed logical to first define the problem. It was then that I noticed a significant omission in the preparation for the grants competition – there had been no commissioned paper on the demographic facts of the problem. As the language of the grants announcement made clear, everyone involved was thoroughly convinced that the problem existed and that it was serious.

Very quickly, my planned research synthesis chapter turned into two chapters, a first chapter that outlined the demographic facts of the problem

(Chipman & Thomas, 1985) and a chapter attempting to synthesize the findings of the research grants (Chipman & Wilson, 1985). I soon uncovered a major surprise: mathematics has been the least sex-typed of college majors! By that, I mean that the representation of women among math majors has been as close to their representation among all recipients of Bachelor of Arts (BA) degrees as one can find for any field of study. This fact immediately casts doubt on the idea that mathematics is a particularly problematic field for women. It was revealed by a readily available and complete data set, the statistics on earned degrees conferred in the United States that have been maintained by the National Center for Education Statistics (NCES) since at least the 1949–1950 academic year. In that academic year, 24% of all BA degrees went to women and nearly 23% of BA degrees in mathematics went to women. In the 1976–1977 academic year, the last year for which statistics were available when I did these analyses, 46% of BA degrees were awarded to women and 42% of the BA degrees in mathematics. In publications over the years, I have periodically updated these figures. My latest update appears in Table 1.1. Note that women’s share of the degrees awarded remains high at the BA level (although lagging their recent majority status among BA recipients) and has continued to climb at the level of graduate degrees. In the early 1980s, I concluded that if there was any problem concerning women’s participation in the study of mathematics, it seemed to be at the level of continuation to the doctoral degree and that some self-examination of university math departments might be warranted. Despite some improvement, this conclusion still seems valid. Women’s level of participation in the study of mathematics itself has been much higher than their level of participation in other fields that are seen as math-related, requiring mathematical competence, such as engineering, computer science, and physics. Thus, it hardly seems plausible that

TABLE 1.1. *Percent of Degrees Awarded to Women*

	BA – All	Math BA	MA – All	Math MA	PhD – All	Math PhD
1950	24	23				
1960	35	27				
1970	43	37				
1975	45	42	45	33	22	10
1980	49	42	49	36	30	13
1985	51	46	50	35	34	15
1990	53	46	53	40	36	18
2000	57	47	58	45	44	25

Source: The primary source of these data is the Series of Earned Degrees Conferred, National Center for Education Statistics. Data 1950–1970, as cited in Chipman & Thomas (1985). Data 1975–1990, as cited in National Science Board, *Science and Engineering Indicators – 1993*, Appendix Tables 2-19, 2-25, and 2-27, pp. 272–285. Data for 2000 from *Digest of Education Statistics* (2001), <http://nces.ed.gov/pubs2002/digest2001/tables/>.

aversion to mathematics is or was functioning as an important barrier to women's participation in those fields. Perhaps the explanation should be sought elsewhere.

No such complete data were maintained concerning the study of mathematics at the high school level. However, at the time of my synthesis effort, I was able to find a number of large representative data sets. One of the grants had been to Armstrong (1985) for a National Assessment of Educational Progress (NAEP) survey of women and mathematics that was conducted in 1978, taking a nationally representative sample of 1,700 twelfth-grade students. Thirty-one percent of the males and 27% of the females had taken some variant of the usual 4-year high school mathematics sequence. Similarly, the 1979 report of the College Entrance Examination Board (CEEB; Educational Testing Service [ETS], 1979) stated that 64% of males and 45% of females expected to have completed four years or more of high school mathematics. Of course, individuals taking the SAT are not a random sample of all students, but they constitute a large fraction of students going on to college. More than 900,000 individuals were covered by that 1979 report.

These data did indicate a sex difference in the study of high school mathematics, especially in the study of advanced courses such as calculus or optional courses beyond the standard college preparatory track: those courses tended to be about 60% male in participation. However, these differences were not nearly so extreme as most people believed or as Sells (1973) had reported. About 40% of those who were approaching college with 4 years of mathematics preparation were women and about 40% of women were entering college well prepared in mathematics, having taken the standard 4 years of high school mathematics. (For more details, see Chipman & Thomas, 1985.)

There were also older data sets that could have better informed the research planning. The National Longitudinal Sample of persons who were twelfth graders in 1972 showed that about 39% of the males and 22% of the females had taken 4 years of high school mathematics. Farther back, the 1960 Project TALENT sample showed that 33% of the boys and only 9% of the girls were taking four years of mathematics. Even so, it would have been difficult to argue that mathematics was functioning as a barrier to entry into math-related careers because only 3% of the girls were planning to go into math-related careers. Clearly, too, a significant change had occurred between 1960 and 1972: the percentage of girls studying 4 years of high school mathematics had more than doubled. The successive CEEB reports from 1973 to 1979 also showed a slow increase in female participation in the study of advanced high school mathematics. It seems that a process of change was well underway by the time the grants competition was initiated. One wonders how the research would have been different if these facts had been recognized at the time. Why weren't these facts recognized? Why weren't such analyses done in preparation for the grants competition?

Perhaps it was that the decision-makers and the lobbyists for the research harkened back to their own school experience in the 1960s, 1950s, and before and remembered that few girls had been studying advanced mathematics in those days or perhaps remembered that they themselves had not chosen to study mathematics. Although an analysis of the dimensions of the “problem” seemed like a mundane, standard thing to do when starting the research synthesis effort, perhaps I asked the question because I myself had majored in mathematics in college and had attended a high school in suburban Chicago where many girls had studied advanced mathematics in the early 1960s. A large social change in expectations for women’s lives occurred during those years; undoubtedly some women found themselves hampered by the educational choices they had made when expecting to lead very different lives. Analyses of the Project TALENT data (Wise, 1985) showed that the choice to study advanced mathematics in high school in 1960 was predicted by a girl’s expectation of going on to college and pursuing a career of some sort. In later years, many more girls would have such expectations. Correspondingly, it seems that by the time the 1998 High School Transcript Study was done, sex differences in high school math course participation had disappeared, or even shifted to favor females. Even calculus was shown as being taken by 11.2% of males and 10.6% of females; Advanced Placement (AP) calculus by 7.3% of males and 6.4% of females (NCES, 2001).

In summary, by the time the brouhaha concerning the mathematics preparation of young women was raised, the “problem” had already diminished significantly, and that trend has continued until the present time. Sells’s highly publicized and influential data were unrepresentative of the national situation at the time; perhaps her sample size was too small or perhaps the University of California was atypical. Furthermore, the bare facts, as well as some of the analyses done in the studies that provided the facts, cast doubt on the assumptions that were held about the causal relations between the study of high school mathematics and entry into fields seen as “math-related.” It might be that the intention to go into a math-related field, or even the mere intention to attend college, “causes” the study of advanced high school mathematics, rather than vice versa.

Despite what these facts show, it is obvious that the belief that there is a large “women and mathematics problem” persists today. One constantly reads of efforts to “solve” it by offering single-sex math classes and the like.

INVESTIGATING THE DETERMINANTS OF MATH COURSE ENROLLMENT AND ACHIEVEMENT

The primary focus of the research grants that NIE awarded was on understanding the factors determining enrollments and achievement in advanced high school mathematics. Beyond that, the emphasis was on examining

variables that might plausibly explain sex differences in math course enrollments. A consequence of that concern was a relative neglect of cognitive variables in the research that was done. Despite the widespread belief that there are sex differences in some inherent ability to learn and do mathematics, a topic to be discussed later in this section, it was already known in 1977 that sex differences in mathematical ability and/or achievement at the beginning of high school were negligible and, therefore, had little promise of explaining the differences in enrollment or choice of occupational field. Measures of spatial ability were well represented in the research, but measures of general intellectual ability, or prior mathematics ability and/or achievement, were not. Affective measures of attitudes related to mathematics, mathematics study, mathematics teachers, and so on, were well represented. As with demographic facts, the effort to synthesize the results of the research studies brought out shortcomings in the way the research studies had been designed to address the question of determinants of course enrollment. The grant to analyze previously collected, nationally representative Project TALENT data (Wise, 1985) revealed that the strongest correlates or predictors of individual differences in advanced mathematics course enrollment were measures of cognitive ability, mathematics ability, or even verbal ability at the beginning of high school, although these measures did not serve to explain the sex differences in enrollment that were still large at that time. The sex differences in enrollment then present did, however, tend to account for the sex differences in mathematics achievement that were measured at the end of high school. This agreed with Fennema's (1974) earlier report that sex differences in math course-taking had an important role in explaining what had tended to be interpreted as sex differences in inherent mathematical ability. Project TALENT was not designed to examine decisions to enroll in advanced mathematics and science or the sex differences in those decisions. Consequently, it did not include measures of attitudes toward mathematics, and did not provide an opportunity to assess the relative explanatory contributions of cognitive and affective variables. This proved to be a problem for the research program as a whole.

The nature of this problem was evident even within the cognitive realm. As mentioned above, the best-represented cognitive variables in this research were various measures of spatial ability. It was believed that there were sex differences in spatial ability and that spatial ability was important to mathematics. Intuitively, the capacity to mentally rotate, translate, and transform objects appears to be important in mathematical thinking, at least in geometry. Fennema (1977) reported on opinions from the mathematical community that support this point of view.

There is a tendency to think that measures of ability have a stronger theoretical, scientific basis than they actually do. Ability testing and the definition of abilities has been a pragmatic and empirical technology.

Performance is sampled within a domain of tasks or situations, in a way limited by the practical constraints on testing. Statistical techniques, usually factor analysis, are used to identify tasks that “go together,” have something in common, and those that seem to be independent of each other. The hypothetical “something” in common is called a factor, and may sometimes be labeled an ability, although the technical psychometric use of the term *ability* does not always carry with it all the implications of the popular meaning of *ability*. For instance, a psychometric *ability* sometimes consists entirely of learned knowledge. In the history of cognitive testing, it has been found that all intellectual performances have something in common: that is, persons who do well or poorly on one intellectual task also tend to do well or poorly on other, quite different intellectual tasks. This common factor has been called general intelligence or “g.” Some relatively recent research is beginning to show the way to a deeper theory about the nature of general ability. For example, Carpenter, Just, and Shell (1990) showed, by constructing computational models of cognitive processes in solving Raven Progressive Matrices items, and by converging evidence from another task, that individual differences in performance on this well-accepted measure of general intelligence are largely accounted for by individual differences in the number of problem-solving goals that can be managed in working memory. Intuitively, this characterization of general intelligence also sounds much like the essence of mathematical ability, as distinct from learned mathematical knowledge.

Many different tasks, which can be performed with diverse mental strategies, have been called tests of spatial abilities. Various tests of so-called spatial abilities do not necessarily have high correlations with each other, as contrasted with their correlations with other kinds of tasks (Lohman, 1979; McGee, 1979). There is no single, unitary spatial ability that these tests are measuring. Lohman (1979, 1988, 1996) concluded that a considerable proportion of performance on spatial tests, especially complex spatial tests, is explained by variation in measures of general intelligence, what all tests of intellectual performance have in common. One of the surprises of the effort to synthesize the results of the NIE grants (Chipman & Wilson, 1985) was that the studies including measures of spatial ability did not provide any strong evidence for sex differences in spatial abilities, despite a previous review concluding that this was a reliable cognitive sex difference (Maccoby & Jacklin, 1974). The nationally representative and relatively large Armstrong (1979) study even reported a statistically significant advantage for 13-year-old females on 15 items taken from the Paper Form Board test. These unexpected results might be due to the tests used (most often the DAT Spatial Relations test, which requires the examinee to select the three-dimensional (3-D) shape that will be formed by folding a two-dimensional (2-D) shape along indicated fold lines), or due to changes over time affecting the experiential influence on “ability” measures, or

due to relatively small sample sizes in many of the studies. Psychometric studies have often had huge sample sizes that make almost any observed difference statistically significant, even though it may be too small to be considered practically significant. At the time, the research studies that had shown substantial sex differences in a spatial ability (Sanders, Soares, & D'Aquila, 1982; Vandenberg & Kuse, 1979) used a test involving the rotation of objects in three-dimensional space. Indeed, a formal and thorough meta-analysis of the research on sex differences in spatial abilities done independently at about the same time (Linn & Petersen, 1985) concluded that sex differences are found primarily on that type of measure and not on the other types of measures of spatial ability. Although that review has been cited more than 400 times in the intervening years, none of the citing articles is a later review or meta-analysis that would change this picture.

Despite these results undermining the notion that putative sex differences in spatial ability might explain putative sex differences in math enrollments or achievement, it is probably worth mentioning that the evidence for a specific contribution of spatial ability to mathematics performance, distinct from the contribution of general intelligence, is surprisingly weak. Smith (1964) and Werdelin (1961) are two of the most frequently cited references on this point, but neither of them actually provides strong evidence for a relationship between spatial ability and mathematics performance. Several reviewers of the literature have concluded that no such relationship has been shown (Fruchter, 1954; Very, 1967; even for geometry: Werdelin, 1961; Lim, 1963). Fennema & Sherman (1977, 1978) and Sherman (1980) did report that the DAT Spatial Relations test shows a correlation of about 0.50 between the DAT score and general tests of mathematical achievement in a high school population enrolled in college preparatory mathematics courses. However, the DAT is the type of spatial ability test that Lohman (1979) characterized as being similar to measures of general intelligence, and Fennema and Sherman do not provide any evidence for a specific unique contribution of spatial ability either. In the larger and more broadly representative Project TALENT sample, there were two measures of spatial ability, Visualization in 2-D and Visualization in 3-D, but they were not among the variables having a correlation of 0.20 or higher with mathematics achievement (Wise, Steel, & MacDonald, 1979). One of the NIE studies that emphasized spatial ability provided an intriguing pattern of results. Stallings (1985) used the DAT and course-specific tests of mathematics. The pattern of correlations she found for the different types of mathematics is quite consistent with what one might expect: algebra I (0.49), geometry (0.53), algebra II (0.15), trigonometry (0.38), analytic geometry (0.68), and calculus (0.20). Unfortunately, the design of her study did not include a measure of general intelligence or even one of verbal ability, so it, too, cannot provide evidence of a unique contribution of spatial ability to performance in any of these mathematical fields, despite

a suggestion of some promise for analytic geometry and geometry. This seems to have been an opportunity missed because of the intense focus on possible explanations for sex differences. In contrast to the general lack of evidence of a contribution of specifically spatial abilities to mathematics performance, there is such evidence for predictions of success in courses such as mechanical drawing and shop (McGee, 1979).

Within the mathematical community, there is a long-standing distinction between algebraists and geometers. Perhaps this is grounded in a difference in their reliance on spatial thinking, but both types are counted as mathematicians. There is more than one way to do mathematics. The need to do mental rotation in depth (apparently the primary locus of sex differences in spatial ability) probably does not arise all that often. Furthermore, very advanced mathematics often deals with N dimensions, not just 3. Heavy reliance on spatial thinking can prove a barrier in moving on to N dimensions.

AFFECTIVE VARIABLES

In addition to spatial ability, the NIE studies emphasized the possible role of affective variables in determining course enrollments and mathematics achievement. Fennema and Sherman (1976) developed a thorough and extensive set of attitude scales, but two variables have received the most extensive exploration: liking for mathematics and mathematics anxiety/confidence. Although these variables seem closely related conceptually and have a strong correlation with each other (0.60–0.65), they behave rather differently with respect to sex differences (Chipman & Wilson, 1985). Consistently, there is no sex difference in liking for mathematics. Thus, it may not be surprising, after all, that women have been so well represented among math majors. In contrast, there is an equally consistent sex difference in mathematics anxiety/confidence (Hyde, Fennema, Ryan, Frost, & Hopp, 1990). Although Fennema and Sherman (1977) attempted to construct separate scales for anxiety and for confidence, the two scales were found to have a correlation of -0.89 with each other, so they can be considered to have been measuring the same thing. It is not entirely clear what to make of the small mean sex differences that are observed. Because no one seems to have published the full *distributions* of male and female scores, it is not clear, for example, whether serious mathematics anxiety is more common among females than among males. It might be that, for social reasons, females are less willing to express high confidence in themselves as learners of mathematics, even if they in fact have such high confidence. The interpretation of these attitudinal variables is not entirely straightforward. The questionnaires that measure these variables are fallible yardsticks. Some people will use extreme values on the scales; others will not. The expression of true opinions may be tempered by the person's impression of what is a

socially acceptable answer. Expressions of very high confidence in mathematical ability may be more socially acceptable in males than in females. Admissions of weakness, anxiety, or distress may be less socially acceptable for males. We can never be certain that the apparent sex differences in these subjective variables reflect genuine differences in the characteristic the scale purports to measure. Nevertheless, sex differences in mathematics anxiety/confidence showed some potential to explain sex differences in enrollment.

Another important affective variable is the perceived utility of mathematics study and of the resulting mathematical knowledge. Looking over the historical changes in girls' and women's study of high school mathematics, participation in higher education, and participation in the workforce over the past 50 years, it seems likely that the primary driver of change lay in this area. Among the NIE grant studies (see Chipman & Wilson, 1985, for details), the general perceived usefulness of mathematics was moderately related to enrollments, while more specific perceptions of mathematical requirements for a planned job or career or aspirations for higher education had a somewhat stronger relationship to enrollments or enrollment intentions. Wise (1985) reported that sex differences in career interests in the Project TALENT sample from the early 1960s predicted math course enrollments, preceded differences in achievement, and probably could explain the sex differences in enrollment and achievement that then prevailed. As discussed earlier, by 1998–2000, sex differences in high school mathematics enrollment had virtually disappeared and women had become the majority among BA recipients. Yet, sex differences in participation in the so-called math-related fields, engineering (23% female in 2000), physics (22% female in 2000), and computer science (28% female in 2000) remain substantial (NCES, 2001). Other sciences such as biology (58% female in 2000) and chemistry (46% female in 2000) now have an excellent representation of women. The once male-dominated fields of medicine (6% female in 1960; 43% in 2000) and law (2.5% female in 1960; 46% in 2000) changed radically between 1960, the year that Project TALENT began, and 2000. For many women, the primary utility of math study in high school may be in meeting the requirements for admission to the college of their choice rather than the inherent requirements of their occupational choice.

Thus, the historical evidence strongly suggests that the utility of mathematics study for girls and women was an important factor in changing rates of participation in advanced high school mathematics courses. However, one of the frustrations in summing up the results of the NIE math grants and similar research done at that time was the difficulty in performing analyses that would shed light on the relative importance of various cognitive, affective, and other variables in predicting mathematics enrollments, intentions to enroll, and mathematics achievement. Not surprisingly, earlier mathematics achievement, confidence in oneself as a learner of mathematics, and

liking for mathematics – among other variables studied – are all related to each other. Therefore, it is difficult to say which of these variables should be regarded as more basic or “causal,” to which should be attributed the effects of what these variables have in common. These data resisted efforts to define approximately independent but still meaningful variables that would make the results of regression analyses reliable and meaningful. Overall, the results suggested that general cognitive ability was the most important variable predicting individual differences in mathematics enrollments, and that mathematics confidence/anxiety probably made some independent contribution. The evidence for the independent contribution of perceived utility of mathematics was even weaker, largely because so few studies included a good representation of both cognitive and affective variables. Also, a number of studies had data only about reported intentions to enroll in advanced math courses, not about actual enrollments. The latter was a harder, more predictable variable. The NIE studies also included many efforts to measure social influences on student enrollment and achievement, including the behavior, attitudes, or perceived attitudes of parents, teachers, and peers. However, none of these variables emerged as important, and sex differences were not usually found. Disappointingly few strong conclusions could be drawn from the research because the strong focus on sex differences resulted in a poor representation of cognitive variables. Ignoring the variables with the strongest relations to the predicted variables made it difficult to gauge accurately the size of the influence of the affective variables, given the strong intercorrelations.

BARRIERS TO WOMEN’S PARTICIPATION IN MATH-RELATED CAREERS – THE BIG PICTURE

Having raised the opposite causal possibility that career expectations may influence math course participation, let us turn to consider the notion that adequacy of mathematics participation is or was functioning as a barrier to women’s entry into mathematics-related careers. Dunteman, Wisenbaker, and Taylor (1979) performed analyses of the National Longitudinal Survey (circa 1972) data that attempted to understand sex and race differences in the selection of engineering and science majors. Their model for predicting the selection of a science major had four variables on which females obtained lower scores: orientation toward things ($2/3$ standard deviation [SD] lower) vs. orientation toward people, reported mother’s educational aspirations for the child ($3/10$ SD lower), mathematics score ($1/4$ SD lower), and number of science courses taken (considerably lower). Even after these variables having a negative effect for females were taken into account, an unexplained direct negative effect of being female upon selection of college science remained, and it was about twice as important as the math score. Additional variables considered to be related to women’s roles or values

also failed to explain that difference. It would be interesting to know if this unexplained barrier to women's entry into those fields still exists. If so, that should probably be regarded as a social problem needing to be addressed. The large sex differences in vocational interest patterns, such as the crude dimension of interest in things vs. interest in people that Dunteman et al. (1979) constructed are probably the most important single factor in explaining the low representation of women in what have been thought of as math-related careers. It is surprising that so little research attention has been given to this explanation for sex differences in career participation. More recent research indicates that these sex differences in vocational interest patterns continue to exist (Hansen, 1988; Lippa, 1998). Unfortunately, little is known about how such vocational interests develop, except that they seem to develop rather early in life (Tyler, 1964). A citation search on Tyler's article did not reveal recent developmental research of this kind. It is not obvious to me, however, that sex differences in the area of interests, or the resulting differences in occupational choices, should be regarded as a problem.

MATHEMATICS ANXIETY AS A BARRIER

An opportunity for me to pursue the open question about the possible influence of mathematics anxiety/confidence on the selection of a field of study, separate and distinct from the influence of mathematics ability/achievement, arose in a study of three successive classes at Barnard College (a women's college associated with Columbia, then and now a coeducational college but once a men's college), totaling about 1,360 women initially and 1,074 for whom complete data through college completion were available (Chipman, Krantz, & Silver, 1992, 1995). This study was able to demonstrate a strong influence of mathematics anxiety/confidence, independent of the effects of quantitative SAT scores. In this select population (mean QSAT about 600), QSAT had no effect on expressed interest in a scientific career at the time of college entrance, whereas mathematics anxiety/confidence did have a significant effect. For actual biological science majors, the same picture held true at the end of college: no effect of QSAT, significant effect of mathematics anxiety/confidence. For actual physical science majors at the end of college, both QSAT and mathematics anxiety/confidence showed significant effects on the likelihood of a major. In this population, there were some individuals with very high QSAT scores and low mathematics confidence. The results of this study indicate that the sex difference in mathematics confidence may be partially responsible for some of the underrepresentation of women in science and engineering fields, but the sex difference in math confidence seems to be quite small among able college students (Hyde et al., 1990). In the Barnard study, the full impact of that effect was expressed *prior to* college entrance, but it was

substantial: the odds of becoming a science major were 5 times as great for the math confident as for the math anxious. The effect of QSAT on actual completion of a physical science major was even more substantial – the odds of being a physical science major were 16 times as great for the group with the highest QSAT (scores over 650) as for the lowest group. There are several possible interpretations of this QSAT result. It may reflect a reality that mathematical competence is important in the pursuit of a physical science field. Alternatively, it may reflect a strong belief in the college-level community – both faculty and students – that mathematical competence really matters. Students with less than the highest QSAT scores may be counseled out of physical science fields or may counsel themselves out of those fields. When interpreting these results, however, it is important to remember that the Barnard College population was an intellectually select population. According to National Science Foundation (NSF) data, the mean QSAT of the Barnard population was as high or higher than the mean QSAT of all U.S. males receiving BA degrees in physical sciences, engineering, or even mathematics at that time (NSF, 1986). The majority of those Barnard students, therefore, were probably capable of completing a physical science major. However, interest in a science career was rare at Barnard College – there were 45 physical science majors, 69 biological science majors, 572 social science majors, 357 humanities majors, 31 in creative writing or similar fields, and no mathematics majors at all. Although interest in a science major at the beginning of college was strongly predictive of an actual major, such interest was so rare in this population that there were actually more physical science majors coming from the group classified as not open to consideration of engineering or science careers at the time of college entrance than from those with high initial interest. Contrary to popular belief and some prior research results, it seems that experiences during college can result in a science major. The *people vs. things* dimension of vocational interests was also investigated in the Barnard study; its influence on occupational interests and major selections was substantial and had had its effect prior to college entrance.

MATHEMATICS ABILITY AND ACHIEVEMENT AS A POSSIBLE BARRIER

Finally, let us turn to the subject of possible sex differences in mathematical ability and/or achievement. As was pointed out earlier for the case of spatial abilities, there is no deep theory about the fundamental nature of mathematical ability. It has proved difficult to define mathematical ability factors that are any more predictive of success in mathematics than measures of general intelligence (Aiken, 1973). As noted above, Carpenter, Just, & Shell's (1990) characterization of the fundamental basis of general intelligence is also very plausible as a characterization of the essence

of mathematical ability. Beyond general intellectual ability, the cognitive variable that predicts future mathematics performance is past mathematics achievement. Previous grades in mathematics appear to be the best available predictor of success in college mathematics (Wick, 1975).

In the United States and elsewhere, there is widespread belief that males outperform females in mathematics. However, the data for the United States do not necessarily support this belief (Chipman & Thomas, 1985; Hyde, Fennema, & Lamon, 1990). Large, representative studies of U.S. student populations have tended to find little or no sex difference in overall mathematics performance prior to the secondary school years, when the study of mathematics often becomes optional in the United States. Despite occasional reports that boys or girls in elementary school perform better on one or another type of math test item, a meta-analysis concluded that no such differences are evident prior to secondary school (Hyde et al., 1990). An exception to the general picture of equality is that searches for mathematically talented youths have generated reports that extremely high levels of mathematical performance on the SAT at a young age (about 7th or 8th grade) are much more frequently found in males (Benbow & Stanley, 1980, 1983). These reports have received much publicity and have had a substantial effect on public beliefs about male and female performance in mathematics (Eccles & Jacobs, 1986). Because of the way in which these searches are conducted, methods that do not ensure representative sampling, it is difficult to know what one should conclude about the actual incidence of high levels of mathematics performance among young male and female students in the United States. The U.S. National Assessment of Educational Progress (NAEP) (National Science Board, 1993), which does aim at achieving a nationally representative sample, reported that 0.2% of females at age 13 and 0.5% of males at age 13 attained the highest category of mathematical proficiency (p. 232) – characterized as involving multi-step problem-solving and algebra, as well as various other mathematical content usually taught during high school. However, it is clear that the generalization from these reports to beliefs about the performance of more typical male and female students is not justified. In the general population, sex differences in mathematics performance prior to secondary school are negligible.

By the end of secondary school, however, sex differences in mathematics test performance that favor males have usually been reported in the United States, and the performance differences seem to arise from problem-solving tests or items (Hyde et al., 1990). For many years, the possibility that differences in course taking might account for these differences in mathematics test performance seems to have been ignored. Fennema (1974) pointed this out. Obvious as that hypothesis might seem, many seem to have concluded that such test results implied lesser mathematical ability among female students. Analysis of data from the representative survey sample

of U.S. students that was collected in the early 1960s – when there were substantial sex differences in secondary school mathematics course enrollments – showed that course enrollments statistically accounted for nearly all the sex difference in mathematics performance at the end of secondary school (Wise, 1985). However, as noted above, sex differences in course enrollments diminished greatly over time. Nevertheless, a performance difference on the SAT remained circa 1993. It still remains today. The CEEB (2000) report on college bound seniors, 2000, reported no difference between males and females in number of math courses taken (3.8) but a mean difference of 35 points on the SAT math test, favoring males. Similarly the NAEP (National Science Board, 1993) results showed that 5.6% of 17-year-old females but 8.8% of 17-year-old males were attaining the highest category of proficiency in 1990 (p. 233). On the other hand, the mean results for 17-year-old males and females in 1990 showed no difference (p. 7 & 231), the culmination of a gap-closing trend.

The picture is further complicated by the fact that different measures of mathematical performance yield different messages about sex differences. In an important review paper, Kimball (1989) showed that course performance measures consistently favor females. Similarly, the results of some studies tend to indicate that examinations that are closely tied to the instructed curriculum, like the New York State Regents Exam (Felson & Trudeau, 1991) or IEA Math content, which is well represented in the “implemented curriculum” (Hanna, 1989), are more likely to favor females. This side of the story is further reinforced by another large-scale study drawing from a huge sample of students gathered to investigate the validity of the SAT exam. When males and females were matched by university math course taken (in the same institution) and by performance grade received in the course, it was found that the females had received scores nearly 50 points lower on the SAT exam (Wainer & Steinberg, 1992). In other words, the SAT underpredicted the performance of females relative to males in these mathematics courses, advanced as well as introductory. Reportedly, an earlier internal study at the Massachusetts Institute of Technology (MIT) had similar results showing that the SAT underpredicted course performance at MIT, but I was never able to obtain a report of that study. Such results suggest that the observed sex differences in performance on tests like the SAT may reflect differences in responding to the testing situation itself (Becker, 1990), or that they may arise from extracurricular differences in experience that are related to the content of some such tests.

POSSIBLE SOURCES OF SEX DIFFERENCES IN TEST PERFORMANCE

There are a variety of factors that might contribute to persisting sex differences in math test performance. Significant sex differences remained circa 1990 for course enrollments in “extra” mathematics courses like statistics

and probability that are not part of the standard college preparatory curriculum. Large sex differences existed in enrollments in courses like high school physics that may provide considerable practice in solving mathematics problems. If the majority of secondary school females are taking high school mathematics simply to fulfill college or university entrance requirements rather than to prepare for further study and careers that intrinsically require mathematical competence, their degree of involvement with the subject matter may be less than is common for males taking the same courses.

Studies that have attempted to analyze sex differences in performance on the SAT or similar tests have not yielded any great insight (Chipman, 1988b). Individual test items can be found that show very large sex differences, but the reasons for those differences are not obvious. There has been little or no consistency in the apparent nature of such items from one study to the next. Occasionally, these analyses have appeared to confirm hypotheses that sex differences might be concentrated in items with geometric or spatial content, but this has not proved consistently true. (Furthermore, as discussed above, the belief in sex differences in spatial ability is also rather weakly supported by the actual evidence.) Many have hypothesized that the stereotypically masculine content of mathematics word problems might account for some of the sex differences in performance on such items. Subjective examination of items that do and do not show large sex differences does not provide obvious support for that view. Furthermore, an experimental study of this question that had high statistical power (Chipman, Marshall, & Scott, 1991) did *not* confirm the popular hypothesis that sex-stereotyped content of math word problems would affect performance. A more recent study confirmed this negative finding (Walsh, Hickey, & Duffy, 1999). In several studies of the SAT, it was found that a class of items ("data sufficiency items") that ask whether sufficient data are available to answer the question did consistently favor females (Donlon, 1973; Strassberg-Rosenberg & Donlon, 1975). It is rather hard to say why such items should have favored females: one can speculate that perhaps females tended to actually attempt problem solutions, thereby improving their ability to answer these items correctly while consuming extra time that may have hindered their performance on the rest of the examination. In any case, the historical fact that the Educational Testing Service chose to drop a class of items that consistently favored females from a test that consistently favors males should cast doubt on the tendency to treat the SAT as if it were some gold standard of mathematical ability. (Ostensibly, these items were removed for being too susceptible to coaching.) The SAT is a speeded multiple-choice test that rewards test-taking strategies such as guessing based on partial information. The characteristics that produce good performance on the SAT are undoubtedly somewhat different than the characteristics that result in good performance in mathematics courses