DOCUMENT RESUME

ED 353 270 TM 018 973

AUTHOR Friedman, Lynn
TITLE A Meta-Analysis of Correlations of Spatial and Mathematical Tasks.
PUB DATE 92
NOTE 85p.
PUB TYPE Information Analyses (070) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Correlation; Elementary Secondary Education; Geometry; Mathematics Achievement; *Mathematics Skills; *Meta Analysis; Orientation; *Research Reports; *Sex Differences; *Spatial Ability; Visualization

ABSTRACT

The meta analysis reported in this paper considers the implications of combined correlational evidence for the nature of the relationships of mathematical and spatial skills, and for the possibility that spatial skill underlies gender differences in favor of males on mathematical tasks. In all, 136 studies reported in 116 articles and dissertations were considered. Results indicate that corrected space-mathematics correlations are not high. Geometry-space correlations are surprisingly low. Orientation correlations are generally lower than visualization correlations. Verbal-mathematics correlations are usually numerically higher than space-mathematics correlations; same-study differences are frequently significant, especially for two-dimensional spatial tasks. Thus, correlational evidence does not indicate that spatial skill plays a special role in mathematical achievement as mathematics is taught and tested today. Gender patterns sometimes vary. Females' correlations do not differ according to the cognitive level of the mathematics task as often as do those of males. Females have higher verbal-mathematics than space-mathematics correlations in more categories than do males. Correlational evidence does not support the hypothesis that spatial skill underlies gender differences in mathematics. Seven tables and four figures present data from the analysis. There is a 158-item list of references. An appendix contains 16 additional tables of study data. (SLD)

Abstract

A Meta-Analysis of Correlations of Spatial and Mathematical Tasks

Lynn Friedman, University of Chicago

Reports of substantial correlations between spatial and mathematical tasks have been common in the last half of the twentieth century. Sherman (1967) has suggested that inferior spatial skill is the explanation for females' inferior performance in mathematics. Many studies have reported relevant correlations: These are uniform neither in absolute size nor in size relative to other correlations calculated in the same studies. The meta-analysis reported in this paper considers the implications of combined correlational evidence for the nature of the relationship of mathematical and spatial skills, and for the possibility that spatial skill underlies gender differences in favor of males on mathematical tasks.

Results indicate that corrected space-math correlations are not high. They range from means of approximately .35 for two-dimensional orientation tasks to .47 for three-dimensional visualization tasks. Geometry-space correlations are surprisingly low. Orientation correlations are generally lower than visualization correlations. Verbal-math correlations are usually numerically higher than space-math correlations: same-study differences are frequently significant, especially for two-dimensional spatial tasks. Thus correlational evidence does not indicate that spatial skill plays a special role in mathematical achievement as mathematics is taught and tested today.

Gender patterns sometimes vary: 1) Females' correlations do not differ according to the cognitive level of the mathematics task as often as do males; 2) Females have higher verbal-math than space-math correlations in more categories than do males. Correlations of spatial tasks with mathematical tasks on which gender differences are found are high only for SAT-Q-space correlations. SAT-Q-space correlations also exhibit gender differences: Females' SAT-Q-space correlations are higher than males'. Intervening variables which might explain the high SAT-Q-space correlations and the gender difference are confidence and clusters of interests. Few other gender differences were found. Correlational evidence does not support Sherman's hypothesis that spatial skill underlies gender differences in mathematics.

A Meta-Analysis of Correlations of Spatial and Mathematical Tasks

The resurgence of feminism in the 1960's has brought a renewed interest in cognitive gender differences. In particular, gender differences in mathematical tasks have become the focus of a steadily growing body of research. Attempts to explain these gender differences have led to a scrutiny of the particular types of abstract reasoning involved in the varied tasks of mathematics. This, in turn, has led to an examination of theories held by psychometricians who distinguish between verbal and nonverbal intelligence. Some of these theories posit an essential connection between spatial skill and advanced mathematical achievement.

Many mathematical tasks have obvious spatial prerequisites. Being able to visualize a cube is helpful, when no cube is around, in calculating its surface area. Visualizing the rotation of a right triangle about the extension of one of its legs is surely vital in understanding the calculation of the volume of the generated solid. Mathematical tasks involving the measurement of physical objects or their motions could hardly be accomplished without visualization. However, the spatial skills involved in such tasks are often elementary.

Other mathematical processes do not require spatial imagery in any obvious way. Forming number and operation concepts, sequencing, and solving some of the more difficult types of word problems do not appear to require visualization. Yet researchers have sometimes found that these tasks correlate as well with spatial tests as do more obviously spatial mathematical tasks.

Hypotheses regarding the nature of the connection between spatial and mathematical skills have very practical implications for mathematics educators. If spatial skills underlie mathematical skills, should we not be training spatial skills to increase our students' mathematical competence? And, if so, how should we go about it? Should we train spatial skills separately, or should we give more time and coverage to the mathematical problems with spatial content? Mathematics educators have few doubts about the overall value of spatial skill, but they are unsure of where it plays a role and how it can be developed. They now encourage teachers to develop spatial sense in their students, using a somewhat eclectic "bag of tricks". Should more be done?

Smith (1964) is representative of a group of theorists who believe that spatial and mathematical skills are both outcomes of the same process, or set of processes, of thought. He writes that "...there is a growing awareness that mathematics is primarily concerned with spatial, geometrical or configurational concepts" (page 134). In his view, spatial ability, or something underlying it, will enable those who possess it to reason differently, more effectively. He and others believe that this ability is innate, but may go undeveloped because its wider effects are not understood.

Smith, a factor analyst, has supported his views with correlational evidence. However, a wide range of studies have reported relevant correlations: These correlations are uniform neither in absolute size nor in size relative to other correlations calculated in the same studies. The meta-analysis reported in this paper considers the implications of combined correlational evidence for the nature of the relationship of mathematical and spatial skills, and for the possibility that spatial skill underlies gender differences in favor of males on mathematical tasks.

Correlations of tested skills may be moderate or high because the skills are related. On the other hand, intervening variables such as general intelligence or common interests may cause spurious correlation. Correlations may be low because the skills are unrelated. Another, sometimes overlooked, reason for low correlations is uneven development of abilities: one of the two abilities measured may have been trained while the opportunity to learn the other has been limited or absent completely. For example, in Project Talent data (see Flanagan et al., 1964)., correlations between arithmetic reasoning and advanced mathematics are very low for 9th graders: for 12th graders, they exceed .50.

These four general explanations, two for low correlations and two for high, form the basis of the interpretation of results of this meta-analysis. Conclusions will be presented after a review of the literature, a review of the studies collected for the meta-analysis, a discussion of methodology, and presentation of results.

Review and discussion of the literature on the relationship of spatial and mathematical tasks

Investigations of spatial ability: Sir Francis Galton's (1918) research on visual imagery is frequently cited as the inspiration for modern studies of spatial ability. Among Galton's conclusions were "the visualizing faculty is a natural gift, and like all natural gifts, has a tendency to be inherited..." (p. 69). At the same time he believed that the faculty could be developed by education. Women, he claimed, have higher visualizing power than men, and, moreover, "scientific men, as a class, have feeble powers of visual representation" (p. 60).

Galton argued that this deficit in scientific men occurred because a propensity to create detailed mental images works against the development of "habits of highly-generalized and abstract thought" needed for science; disuse causes the loss of the former skill. However, Galton did not believe this had to be the case. He valued the ability to transform images and use those transformations in problem solving: "this free action of a vivid visualizing faculty is of much importance in connection with the higher processes of generalized thought, though it is commonly put to no such purpose..." (p. 76). In these words lie the seeds of theories that were to be developed later by Smith, Herman Witkin, Raymond Cattell, and others.

Spearman considered spatial ability to be simply one of many particular manifestations of general ability (1927). Yet, as the predictive value of tests of mechanical ability became evident, and as non-verbal paper and pencil tests were developed to measure these and other abilities, factor analysts began to affirm the existence of a "group" spatial factor, interpreted as a spatial acumen or intelligence underlying these tests (e.g.,Kelley, 1928; El Koussy, 1935; Thurstone,1938). As large batteries of tests began to be administered to large numbers of subjects, there was a progression from agreement that there was one spatial ability to accord that there might be several. (e.g., Guilford et al., 1951; Thurstone, 1951 as discussed in Smith, 1964). However, the progression was not straightforward. Moreover, terminology was not uniform among theorists: Differing studies found different tests clustered together.

In 1957, Michael, Guilford, Fruchter and Zimmerman published an article describing three classes of spatial abilities that seemed to embody the characteristics described in Thurstone's later work as well as those of the Guilford group. Two of these classes have been connected to mathematical problem solving by various researchers. Many group tests have been constructed to measure them. Most of the tests are variations of the process of identifying two different pictures of the same object. The chain of reasoning connecting the two may involve either visualizing a rotation or similar transformation of the object as a whole or visualizing movement of parts of the object so that it takes a different shape. The first type of reasoning is currently believed to be holistic reasoning, and is usually denoted "spatial orientation" skill; the second is characterized as multi-step reasoning allowing trial and error checking of features of the object, and is usually denoted "spatial visualization" skill.

Schonberger (1976) has given a mathematical characterization of the distinction of orientation from visualization, noting that the former involves simple rigid transformations of whole objects. She hypothesized that three-dimensional transformations were more difficult to visualize than two-dimensional ones, and thus distinguished four categories of spatial tests in her work. We have used her categorization in this work. The dimensional breakdown is important because of the greater familiarity of two-dimensional tasks: Two-dimensional orientation tasks are frequently taught in school; two-dimensional visualization tasks can be found in children's magazines: puzzles using hidden figures are especially popular.

ETS's Card Rotations is a typical test for orientation in two dimensions (o2): a sample item from it is shown below. The respondent is asked to distinguish reflections from rotations of the figure to the left of the line in those figures to the right of it.

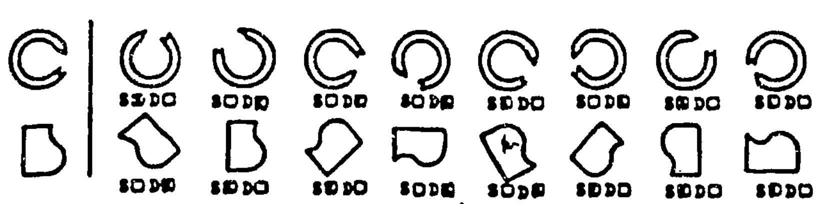

FIG. 1. Sample items from ETS's Card Rotations Test. Reprinted by permission of Educational Testing Service, the copyright owner.

Orientation in three dimensions (o3) is often tested by using the Vandenberg-Shepard Mental Rotations Test. The respondent is asked to identify those figures to the right of the one in the circle which are rotations of it.

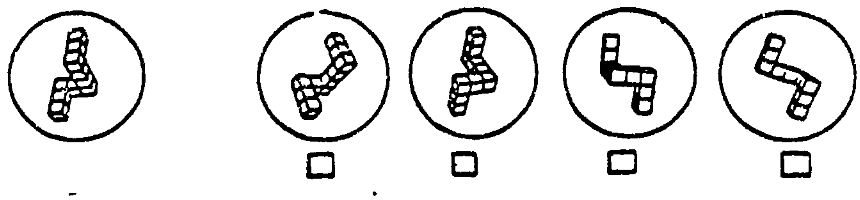

FIG. 2. A sample item from the Vandenberg-Shepard Mental Rotations Test. Used by permission, Steven G. Vandenberg, Institute for Behavioral Genetics, Boulder, Colorado.

The Minnesota Paper Form Board Test is a typical two-dimensional visualization (v2) test: the respondent is asked to identify the configuration in the lettered square that is the result of putting together the shapes in the numbered square:

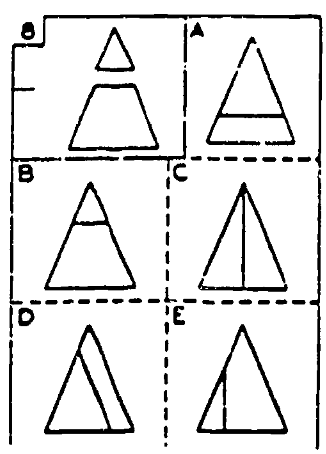

FIG. 3. A sample item from the Minnesota Paper Form Board Test. From the Revised Minnesota Paper Form Board Test. Copyright 1948, 1970 by the Psychological Corporation. Reproduced by permission.

The Differential Aptitude Spatial Relations Test has been used as a test of visualization in three dimensions (v3) in many studies: the respondent is to select the figure on the right that can be obtained by folding the plane figure on the left.

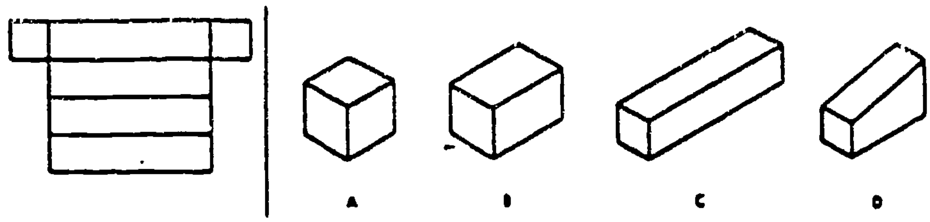

FIG. 4. A sample item from the Differential Aptitude Space Relations Test. From the Differential Aptitude Tests: 4th Edition. Copyright © 1982, 1972 by The Psychological Corporation. Reproduced by permission. All rights reserved.

Both two- and three- dimensional visualization tests can be broken down into sequences of steps involving rigid transformations of two-dimensional figures, sometimes in three-dimensional space. Thus it is conceivable that the sets of tests of o2, v2, and v3 form a series of increasing difficulty. The o3 tests do not seem to fall in this series in any natural way. Guay and McDaniel (1978) and Zimowski and Wothke (1985) argue that these are the best tests of pure spatial reasoning precisely because their items are most resistant to solution by analytic processing, the kind of processing involved in visualization tasks.

<u>Investigations of mathematical ability</u>: The idea of a single mathematical ability is not entirely a lay notion: Binet , Hadamard , and Poincare (1910, 1945, and 1908, respectively, as discussed in Krutetskii, 1976) all thought that there was a mathematical ability that was, in some sense, unitary. Today, Howard Gardner (1983) articulates the notion of a single "logico-mathematical" ability, at the core of which are the abilities to handle long chains of reasoning and to discover analogies between mathematical abstractions. These processes are not necessarily specific to mathematical ability, and Gardner does neglect the specific, very abstract and nonpersonal, content of mathematics.

Plato may have been the first to distinguish types of mathematical content: arithmetic with number as its content, geometry with form, were subjects of the intellect as opposed to subjects of opinion. Since Plato's time, mathematical content has mushroomed. It is the nature of the varied content areas and their applications that has generated doubts about the unity of mathematical reasoning.

Early in this century, analyses of Courtis (1910) and Stone (1910) (see the discussion in Werdelin,1958, p. 53) and of Rogers (1918) found diverse mathematical abilities. Results from factor analysis, a technique used frequently from the 1920's to the 1980's, have been scattered and sometimes contradictory with respect to the number and kinds of mathematical factors found (e.g., Thurstone, 1938; Wilson, 1933; Wrigley, 1958). Barakat (1950 and 1951) carried out a large study of younger students, concluding that verbalization may obstruct the gestalt-like grasp of mathematical notions, hindering mathematical thinking. Guilford and his colleagues carried out many factor analyses of armed services personnel. In these studies, reported in the 1950's, mathematical and verbal reasoning tests often loaded on the same factor.

Observational and factor analytic studies were not the only source of illumination of mathematical ability in the first half of the century: Several introspective studies were published, including two by prominent research mathematicians (Hadamard, 1945; Poincare, 1908). According to Krutetskii, Poincare was the first to distinguish two types of mathematical thought, which today would be characterized as analytic and holistic. Soviet psychologists, who have perceived written tests as instruments of the bourgeoisie, used for class oppression, prefer clinical studies: Krutetskii himself modeled the type of clinical study that is increasing popular with mathematics educators in the United States today, the observation of problem-solving protocols of individual students (see, e.g., Schoenfeld, 1985; Fennema and Tartre, 1984).

Werdelin (1958) gave particular attention to analyses of older or more talented students, in which he found what he dubbed a "mathematical reasoning factor"; it had substantial loadings for geometry tests, arithmetic word problems, algebraic equations, and spatial visualization tasks. This factor was not deductive. Plato considered demonstration, or proof, the primary mode of reasoning for mathematics, as for any science: problem solving was a matter for artisans in the market place or warriors on the battlefield, not for philosophers. Neo-Platonists, those who claim that space and number are the primary content of "core" mathematics and that deduction is its primary form of reasoning, are still with us (e.g., Moore and Witmer, 1991). Yet market places and battlefields are held in increasing respect by many mathematicians and educators today.

In order to analyze the relationship of mathematical and spatial skills, we have used the categories of mathematical tests familiar in standardized testing. "Problem solving" or "application" subtests are common in standardized tests of mathematical achievement today. They, along with tests characterized as tests of mathematical concepts, form the basis for judgments about higher level skills in mathematics. Computational subtests, considered tests of lower level skills, are also the rule. Mathematical deductive skill is rarely tested except in geometry proofs.

<u>The connection of spatial and mathematical skills:</u> Interest in perceptual and performance tests as measures of intelligence increased in the 1930's and 1940's. This led to two major developments in psychology in the United States.

The first was Cattell's articulation (e.g., 1971) of the notions of fluid and crystallized intelligence. Cattell was skeptical that verbal intelligence tests measured innate ability: His doubts were reinforced by the findings that performance on the nonverbal, figural tests seems to stabilize at about 13 years of age whereas performance on verbal tests improves over all the school years. Thus he formulated two concepts, "fluid intelligence" to characterize basic nonverbal skill in reasoning, "crystallized intelligence" to characterize reasoning as embellished by school and cultural learning. Cattell writes that performance tests, which measured fluid intelligence, came to be characterized as "culture fair" tests. However, not everyone agrees that they deserve the label (see, e.g., McFie's study of African youth [1961]).

Another group of psychologists was as influenced by perceptual tests measuring "flexibility of closure" and related skills as Cattell had been by other performance tests. Led by Herman Witkin, they developed the notion of "field independence". Field independence represents a nonverbal ability to discern the salient features of a situation and thus to solve problems and plan tasks effectively. Witkin and others believed that field independence is related to "complex arithmetical tasks" (1962, p. 201). Research on field independence has dwindled since Witkin and Goodenough recognized that different tests of the concept were not always well correlated (1981). The less specific notion of an underlying spatial nature to mathematical skill, as outlined by Smith, remains with us. In the 1980's and 1990's, this notion has become interwoven with theories of brain hemispheric dominance.

Brain hemispheric research has found its way even into the popular press. Left- and right- brain processes have been distinguished, each with its own set of adjectives so firmly embedded in the descriptions that meanings seem to merge. Left-brain processes are variously characterized as multi-step, analytic, verbal-analytic, and prone to feature extraction. Right-brain processes are depicted as holistic, intuitive, gestalt-like, employing parallel (instantaneous) comparisons.

Spatial skill is usually considered a right-brain skill, verbal skill a left-brain skill. As we have seen earlier, spatial skill has been spliced again, into the purely holistic orientation skills and the more analytic, multi-step visualization skills. These dichotomies and trichotomies are not well-defined. As not all spatial skills are considered purely holistic, not all verbal skills are analytic: Vocabulary skills are not analytic; verbal "cloze" procedure skills are very similar to those spatial closure skills entitled "gestalt-completion" skills. However, the distinction of holistic from analytic skills has become important to discussions of mathematical problem solving.

Some mathematics educators suggest that holistic, "right-brain" skills play a key role in mathematical problem solving (e.g., Wheatley, Frankland, Mitchell and Kraft, 1978). Others believe it is the step-by-step analysis of spatial movement that reflects mathematical reasoning (e.g., Moses, 1977).

Well before brain hemispheric research became widely discussed, Smith (1964) claimed that spatial skills were more important than verbal skills to advanced mathematical achievement. He cited evidence from several factor analyses which he thought indicated that verbal and mathematical abilities were independent except for a shared component of general ability. To explore the question in this research, math-verbal and math-space correlations from the same studies have been compared.

Implications of the relationship of spatial and mathematical skills for gender differences. Gender differences in favor of males have frequently been found on both spatial and mathematical tasks (El Koussy, 1955; Maccoby & Jacklin, 1974), though they have been decreasing or disappearing on some spatial tasks (Hilton, 1985; Linn and Petersen, 1985), and nearly vanishing on many mathematical tasks (Friedman, 1989; Hyde, Fennema and Lamon, 1990). Sherman (1967) conjectured that male advantage in spatial ability might be the source of gender differences found in mathematical achievement. She noted that activities that are stereotypically male -- building models and fixing machines, mechanical drawing, map-reading, and sports -- develop spatial skill. While the traditional causal explanation for male advantage is that males are biologically more likely to possess the skills and thus to choose the activities in which the skills are used and further developed, Sherman suggested that sex-role socialization might be a decisive influence in differential spatial practice.

Many educators have found Sherman's hypothesis plausible (e.g., Burnett, Lane, and Dratt, 1979; Ethington and Wolfle, 1984; Fennema,1980). Study reports of markedly different space-math correlations for females and males have added weight to the conjecture. Some of these studies have found the correlations higher for females, others for males. Tartre (1990) has suggested that spatial skill may be more related to all facets of mathematical performance for females than for males.

Sherman's conjecture shares some of the inference of Smith's argument: If possession of advanced mathematical ability implies possession of spatial ability, then females' lack of spatial ability must lead to inferior mathematical ability. However, even if Smith's theory is incorrect, Sherman might be right if mathematical ability implied spatial ability in some of the mathematical tasks on which gender differences are found. Traditionally, the most consistent reports of gender differences in mathematical tasks have been in the areas of problem solving and applications and of geometry and measurement. More recently, studies of college entrance examinations such as the College Entrance Examination Board's Scholastic Aptitude Test (SAT) have found substantial gender differences, particularly in samples of gifted junior high school students. If Sherman is right, then math-space correlations should be high either in some of these areas or overall.

Questions addressed in this study: Four questions directed this meta-analysis of the relationship between spatial and mathematical abilities. First, does the relationship indicate a pervasive spatial character to mathematical thought? Second, is there a substantial relationship between holistic spatial skill and mathematical problem solving? Third, does the space-math relationship differ for males and females? Finally, does it underlie gender differences in mathematical skills?

These general questions led to specific research questions which guided calculations. The first general question was primarily addressed by four research questions: 1) Do any of the four types of spatial reasoning skills -- two- and three-dimensional spatial orientation skills and two- and three-dimensional spatial visualization skills - have substantial combined correlations with all mathematical tasks? 2) Does the cognitive level of the mathematics skill tested influence the size of correlations? 3) Does the age of the test-taker influence the size of correlations? 4) Are correlations of mathematical and spatial tasks higher than correlations of mathematical and verbal tasks, and do age of subject or level of mathematics tasks influence any differences found?

The fourth research question has implications for the relevance of holistic spatial skill to mathematical problem-solving skill. This second general question was explored through an additional research question: 5) Are orientation spatial skills better related to mathematical problem-solving tasks than visualization spatial skills?

The next research question is relevant to Smith's and other theorists' arguments, as well as to investigation of gender differences. 6) Do geometry, problem-solving, or SAT-Q tasks correlate more highly with spatial tasks than other mathematical tasks?

Finally, gender differences were explored by direct calculation of differences in correlations: 7) Do groups of math-space correlations divided according to the four spatial categories exhibit gender differences? Do groups of geometry-, problem-solving-, or SAT-Q-space correlations exhibit gender differences? Theoretical questions on gender differences were further addressed by considering each of the first six questions separately by gender, in order to see if there were differences in patterns.

Studies collected for the meta-analysis

Reports of studies in journal articles, technical reports, and dissertations were collected, using computerized searches of the ERIC and PSYCHINFO data bases and of Dissertation Abstracts Online. "Spatial ability" and "mathematics achievement" were the descriptors used to select the studies. Only studies carried out after 1950 and before the end of 1990 were collected. Only studies using test results to compute correlations were used. The final group used contained 116 articles and dissertations, reporting on 136 independent studies.

The studies collected exhibit trends over time both in subject matter and statistical techniques. Factor analysis dominates the early studies. Many of the factor analyses collected from the 1950's and 1960's used all-male samples, either of armed services personnel or college or college prep students (e.g., French, 1957,1963; Guilford et al, 1951, 1955; Werdelin, 1958). Corrected correlations range from -.10 to .55, generally not as high as might be expected.

Doctoral students at the Catholic University of America produced another set of factor analyses (e.g., Emm, 1959; McCall, 1955; McTaggart, 1959; Ruszel, 1952). These studies found differences in cognitive strengths between females and males: Emm concluded that girls and boys should not be taught in the same way and thus, perhaps, not in the same classroom. On the other hand, a more recent study by Harris and Harris (1973) found no difference in factor patterns for female and male sixth graders. When differences have been found (e.g., McTaggart, 1959; Wormack, 1980), males' abilities tend to yield spatial factors more often and to be more stratified. Fillela's study of Columbian youth is an exception: here the females were more differentiated than the males.

From the mid 1960's through the 1980's large groups of researchers considered individual differences. One group looked for aptitude-treatment interactions (ATI's) (e.g., Adams and McLeod, 1979; Battista, 1981; Behr and Eastman, 1975; J. P. Becker, 1983; Durapau and Carry, 1981; Kiser, 1986; MacGregor, Shapiro and Niemiec, 1988; McLeod and Briggs, 1980). Results from ATI studies have generally been inconclusive, though when a figure matrix test was used to determine spatial ability, Eastman and Carry (1975) found that those high in spatial but low in verbal ability benefited more from a graphical treatment of quadratic inequalities, whereas those high in verbal but low in spatial ability benefited more from an analytic treatment. Apparently treatments should build upon strengths rather than supplement them.

A large body of studies on field independence has accumulated (e.g., Acker, 1967; Bieri, Bradburn, and Galinsky, 1958; Carment, 1988; Gardner, Jackson, and Messick, 1960; Lynchard, 1988; McKay, 1978; Tabler, 1980; Vaidya and Chansky, 1980). Correlations ranged from about .2 to .6 in these studies. However, problems with distinguishing field independence from general intelligence have arisen in the research, as have disagreements of differing tests of the concept.

A spate of relatively recent studies has directly considered the relationship of spatial and mathematical skills (e.g., Battista, Wheatley and Talsma, 1980; Lean and Clements, 1981; Middaugh, 1979). The relationship of spatial skill to mathematical problem solving has been a particularly popular topic (e.g., Landau, 1974; Schonberger, 1976; Wong, 1984). Landau found that students of low spatial ability solved problems more readily if they were provided diagrams, while students of high spatial ability did best with just the instruction (given to all students) to rank problems according to how helpful a diagram would be in their solution. Wong found that providing visual aids helped students solve problems, but, when visual aids were not supplied, students taught to generate their own visual aids often did so and were more successful problem solvers than other students.

Sherman's conjecture motivated many studies. Schonberger (1976) found that the relationship between spatial skill and performance on math problems for which spatial skill was helpful but not necessary was stronger than the relationship between spatial skills and spatial problems, and somewhat stronger for females than males. Fennema and Sherman (1977, 1978) found some support for Sherman's hypothesis. (However, later studies by Fennema and Tartre (1985) and Tartre (1984) showed that males with low spatial and high verbal skills mathematically outperformed males and females with other combinations of spatial and verbal skills.)

Gender difference researchers are particularly interested in spatial training studies because, if Sherman's conjecture is correct, then gender equity in mathematics can be addressed through spatial training. Many spatial training studies produced correlations for the meta-analysis (e.g., Baldwin, 1985; Moses, 1977; Tillotson, 1984): These studies found spatial training successful, but ineffective in improving mathematical skills. Connor and Serbin (1985), who have been involved in several spatial training studies, did separate correlational studies as well. Researchers interested in developing questionnaires intended to measure spatial experience also have been a source of correlations (e.g., Lunneborg and Lunneborg, 1984, 1986).

Burnett, Lane, and Dratt (1979) and Hyde, Geiringer, and Yen (1975) found evidence to support Sherman's hypothesis: for their college age samples, gender differences in mathematical achievement were insignificant when spatial tests were used as covariates. However, in Pattison and Grieve's (1984) sample of select Australian high-school students, using spatial tests as covariates had little effect on gender differences calculated on groups of mathematical problems.

The Burnett, Lane, and Dratt study was one of many using the SAT-Q as the mathematical measure (e.g., B. Becker, 1978; Gallagher, 1987; Johnson, 1984; Weiner, 1984; Wormack, 1984). For the most part, these studies are of gifted or elite college populations, very selected samples. These studies produced correlations which were atypical in more than one way, and more will be said of them later.

Two large national studies with probability samples, High School and Beyond and Project Talent, produced correlations for the meta-analysis (see Ethington and Wolfle, 1984; Flanagan et al., 1964; Shaycoft et al., 1963). The space-math correlations reported in these studies ranged from about .20 to .48. Project Talent tested a large number of cognitive variables; many of the non- spatial measures, including verbal ones, were more strongly related to mathematics achievement than were the spatial tests.

Methodology

Four types of data were coded from the studies: the necessary statistics, the tests used and their classifications, circumstances of the testing, and characteristics of the sample. Correlations of spatial, mathematical, and verbal tasks were collected. Statistics were recorded for the sexes separately when the information was so reported.

Classification of the spatial and verbal tests was carried out by the researcher. Three colleagues helped in the classification of the mathematical tests: items were placed into categories according to whether they required computational, concept, or problem-solving skills. The tests were then classified as either "computational" or "reasoning," Tests classified as computational often included some noncomputational material, though more than 60% of the items were computational, and more than 80% of the items were computational or conceptual. Intercoder agreement on the tests was 85%. The Brownell Problem Solving Test, the California Achievement Test's Mathematics Total, the Cognitive Abilities Test's Quantitative Subtest, Educational Testing Service's Necessary Arithmetic Operations, the Iowa Tests of Basic Skills, Project Talent's Math I, the Scholastic Aptitude Test's Quantitative Subtest, the Stanford Achievement Test's Applications Subtest, and the Test of Academic Progress's Mathematics Subtest are some of the tests categorized as "reasoning" tests. Examples of tests categorized as "computational" are the California Intelligence Test's nonverbal Numerical Quantity Subtest; the Comprehensive Test of Basic Skills' Computation Subtest, the Differential Aptitude Test's Numerical Aptitude Subtest, the Moore-Castori Algebra Test, the School College and Ability Test, the Science Research Associates Mathematics Concepts Subtest, the Stanford Achievement Test's Computation Subtest, and the Wide Range Achievement Test's Arithmetic Subtest. Test categorizations are available in Friedman (1992).

The only circumstance of testing that varied enough to warrant inclusion as a possibly influential study feature was the year of testing. Sample size, average age, minority composition, and nationality were coded. The academic selectivity of the sample was also coded: Studies of disadvantaged students were coded 1; of average school populations, 2; of college preparatory or college students, 3; of gifted or elite college populations, 4. Study codings are available in Friedman (1992).

Studies often supplied a number of space-math correlations even within one of the four spatial categories and for more than one computational or reasoning mathematics task. Median correlations were used when the number available was odd; when it was even, the correlation just above the median was used. Examples of correlations used from the studies appear in Friedman (1992).

Basic methods of calculating weighted averages of correlations, homogeneity statistics, and random effects model means were taken or extrapolated from Hedges and Olkin (1985). A conversion of anova F's to product-moment correlations was made using an equality appearing in Glass, McGaw, and Smith (1981, p. 150).

Study samples were often large, as were numbers of studies in groups; homogeneity was hard to find. Thus the researcher decided to use random effects models for all groups. If a weighted average of correlations (or their *z*-transforms) was homogeneous in its group of studies, it was considered to be the mean of a random effects model with variance zero.

Comparison of different types of correlations was carried out in two ways. First, if the measure to be combined was the study correlations (z-transforms), the difference of random effects model means for correlations of the two different types was calculated, under the assumption that both means were approximately normally distributed, and thus that their difference was as well.

However, comparison by combining simple z-transforms was not the method of choice for this researcher. A large number of study features may affect correlations. Correction for restriction of range was carried out, using regressions calculated on data from studies which used national probability sample or for which a population standard deviation could be found. Yet measurement error is also known to affect correlations, but can be difficult or fruitless to correct for (personal communication, Hedges, June, 1990). While many study features can be coded, others, such as time (and even year) of testing, location of the sample, and attitudes of researchers and testers are often not reported. An oft-repeated experience for this researcher was to remark that, say, a study number series-space correlation was high compared to average math-space correlations, but then to find that other math-space correlations in that same study were high as well.

Thus the difference in z-transforms of correlations from the same study was the preferred study measure when contrasting correlations of different types. Combining differences in correlations is not problematic when study samples can be separated -- e.g., in gender difference calculations. The variance of the difference is a function of the correlation between the two correlations, which, in independent samples, is zero. However, when the sample is the same for both correlations, the correlation is a function of another intermediate, often unreported, correlation. To circumvent this problem, it was assumed that, for any one group of studies, the correlation, c, was the same for all studies. Calculations were then carried out for three possible values of c: -1, 0 and 0.9, "bracketing" the possibilities for c. If the c's at all three different values produced the same conclusion, that conclusion was assumed to be valid. When differing values of c produced different conclusions, reasonable intermediate values of c were checked.

If the difference of random effects model means was the basis of a comparison, it was called an external comparison. If random effects model means of differences of correlations in the same studies was the basis, the comparison was called internal. Often both types of comparisons were made to answer a question: when results differed, internal comparisons were the primary evidence on which conclusions were based. (When differences occurred, the direction (sign) of the difference was usually the same, but the statistical significance of the results varied.)

Correction for restriction of range was carried out using regression techniques. Population standard deviations could be obtained for about one-third of the tests used in correlations. Correlations using these tests were corrected by a formula from Guilford (1965, p. 363). These corrected correlations are here called "basically corrected" correlations. Basically corrected correlations were used in a regression analysis: The equations they generated were then used to correct all the space-math correlations collected for the meta-analysis.[1]

In order to ensure that results were not artefacts of the correction procedure, calculations made with (regression) corrected correlations were repeated either with uncorrected correlations or basically corrected correlations, whichever of the latter groups was appropriate. The significance level of a few results dropped in the repeated calculations: only results replicated in the comparison calculations are reported here. Because of time restraints, correlations of verbal tests with other measures were not corrected for restriction of range, so, in any comparison employing verbal measures, uncorrected correlations were used throughout.

Results

Outcomes for research questions are presented here. Numerical estimates of differences of correlations will reveal which set of correlations is higher in a numerical sense; however, when confidence intervals for the mean difference cover zero, we shall say "numerical estimates suggest that ...". When the intervals do not cover zero, the language will be "these re lts indicate ..." or "these results show that ...". Restricting vocabulary in this fashion will prevent misstating the strength of the results. The standard significance level is 95%; however, when the p-value of a result is between .10 and .05, a 90% significance level is reported. Tables of statistics for all results may be found in Friedman (1992). Only statistics relevant to significant results are reproduced here.

Question one: Do any of the four types of spatial reasoning skills -- two- and three-dimensional spatial orientation skills and two- and three-dimensional spatial visualization skills - have substantial combined correlations with all mathematical tasks? Table 1 displays weighted averages and other statistics calculated for groups of corrected correlations of mixed gender, female, and male samples. Averages and means range from approximately .33 to .47. They generally form an increasing sequence, with the smallest correlations produced by two-dimensional orientation skills, the next smallest by two-dimensional visualization skills, followed by three-dimensional orientation skills and then three-dimensional visualization skills. Female samples are an exception: both two- and three-dimensional visualization skills produce higher correlations than two- and three-dimensional orientation skills.

TABLE 1: Means and other statistics for sets of independent correlations from the four spatial categories.

	r^+	k	H	r^r	$\sigma^2(\xi)$	$\sigma^2(\psi)$	CI
				Mixed Gender Samples			
o2-math	.354	25	38.9	.364	.0008	.0001	(.343,.384)
v2-math	.405	52	281.6	.425	.0276	.0007	(.381,.468)
o3-math	.410	40	292	.428	.0345	.0012	(.372,.481)
v3-math	.467	55	251.7	.449	.0057	.0002	(.426,.471)
				Females			
o2-math	.335	16	28.5	.329	.0015	.0003	(.297,.360)
v2-math	.439	27	80.8	.444	.0161	.0010	(.393,.493)
o3-math	.352	23	99.4	.380	.0292	.0017	(.307,.447)
v3-math	.451	30	179.1	.450	.0114	.0006	(.410,.488)
				Males			
o2-math	.325	22	56.8	.328	.0033	.0004	(.294,.360)
v2-math	.352	36	111	.372	.0132	.0006	(.328,.413)
o3-math	.376	27	124.2	.398	.0236	.0012	(.340,.453)
v3-math	.439	42	234.3	.409	.0116	.0004	(.375,.443)

NOTE: r^+ = weighted average of the correlations; k = the number of correlations estimating r^+; H = the homogeneity statistic for r^+; r^r = mean correlation calculated from the mean of the random effects model estimated for z-transforms of the group; $\sigma^2(\xi)$ = variance of the random effects model for z-transforms; $\sigma^2(\psi)$ = variance of the mean of the random effects model for z-transforms; CI = a 95% confidence interval for the mean of the random effects model.

Table 8 in The appendix gives some of these statistics for uncorrected and basically corrected correlations. Uncorrected correlations follow the same patterns, though means are somewhat lower. Means for basically corrected correlations are higher, reaching .56 for o3-math, and dropping to .47 for v3-math: however, a substantial subgroup of these correlations are from SAT studies, in which corrections were expected to be high because of selectivity of sample: in particular, the three-dimensional orientation correlations in this group are dominated by an atypical group of SAT-Q-space correlations, which will be described later in the paper.

Differences of random effects means of correlations from the four spatial categories are usually not significant. External comparisons show that the two-dimensional orientation mean is different from the others, but internal comparisons differentiate only the two-dimensional orientation from the three-dimensional visualization category. These remarks hold for all three types of samples -- mixed gender, female, and male -- using both corrected and uncorrected correlations.

Question two: Does the cognitive level of the mathematics skill tested influence the size of correlations? The cognitive level (or degree of abstraction) of the mathematics tests does influence the size of correlations. For mixed-gender samples, space-math correlations are higher for reasoning than for computational mathematics tasks in all but the three-dimensional orientation category.

Differences between correlations of computational and reasoning mathematics tasks are not as common with females as with males. For females, they appear in visualization but not in orientation tasks; for males, they appear in all spatial categories.

Because this pattern difference appeared, additional analyses were carried out for single-gender samples grouped by age. If the average age of the sample is less than 14 years, the sample is described as young. Other samples are described as older. Differentiation between space-math correlations of computational and reasoning math tasks does seem to be a function of age. Young females differentiate in two-dimensional visualization-math correlations. In corrected correlations, young males give some indication of differentiating in the three-dimensional spatial categories: however, this trend does not show up in uncorrected correlations. The older samples reproduce the results for samples of all ages.

Table 2 summarizes the results for samples of all ages and for single gender samples divided by age. Tables 9 and 10 in the appendix give the statistics for these results using corrected correlations.

TABLE 2: Summary of differences in math-space correlations for the four spatial categories and computational and reasoning mathematics tasks.

	Mixed-gender	Females	Males
		Samples of all ages	
o2-math	Reasoning larger	No difference	Reasoning larger
v2-math	Reasoning larger	Reasoning larger	Reasoning larger
o3-math	No difference	No difference	Reasoning larger
v3-math	Reasoning larger	Reasoning larger	Reasoning larger
		Young samples	
o2-math		No difference	No difference
v2-math		Reasoning larger	No difference
o3-math		No difference	Reasoning larger, 90%
v3-math		No difference	Reasoning larger, 90%
		Older samples	
o2-math		No difference	Reasoning larger
v2-math		Reasoning larger	Reasoning larger
o3-math		No difference	Reasoning larger
v3-math		Reasoning larger	Reasoning larger

Question three: Does the age of the test-taker influence the size of correlations? Very few significant differences were found, and these seemed to follow no particular pattern. In mixed gender samples, older students have higher o3-math correlations than do younger students: however, when the one study in this group using a low-level mathematics test is removed, the difference between older and younger students becomes insignificant. Young male samples have higher o2-math correlations than do older ones; older females have higher o3-math correlations than do younger ones.

Question four: Are correlations of mathematical and spatial tasks higher than correlations of mathematical and verbal tasks, and do age of subject or level of mathematics tasks influence any differences found? Verbal-math correlations are higher than two-dimensional space correlations for all types of samples. Three-dimensional space-math correlations are not different from verbal-math correlations in mixed gender samples. However, for single gender samples, verbal-math correlations are also higher than three-dimensional orientation-math correlations. For samples of any gender composition, verbal-math correlations are not significantly different from three-dimensional visualization correlations. However, all numerical estimates of verbal-math correlations are higher than those of space-math correlations.

Division of correlations by age of sample and by level of mathematics test produced mixed patterns. Females often had higher verbal-math than space-math correlations. Young males showed only one significant difference: verbal-math correlations were higher in the two-dimensional spatial category on mathematical reasoning tests. Older males produced more differences in favor of verbal-math correlations: in particular, verbal-math correlations were higher in both two-dimensional spatial categories for mathematical reasoning tests.

Table 3 presents a summary of the results of all comparisons made between verbal-math and spatial math correlations. When differences were significant at the 90% rather than the 95% level, it is noted in the table.

TABLE 3. Summary of verbal-math and spatial-math correlations.

	Mixed gender	Females	Males
		All correlations	
o2-math	Verbal higher	Verbal higher	Verbal higher
v2-math	Verbal higher	Verbal higher	Verbal higher
o3-math	No difference	Verbal higher	Verbal higher
v3-math	No difference	No difference	No difference
		Young samples	
o2-comp. math		No difference	No difference
o2-reas. math		Verbal higher (90%)	No difference
v2-comp. math		Verbal higher	No difference
v2-reas. math		Verbal higher	Verbal higher
o3-comp. math		Verbal higher	No difference
o3-reas. math		Verbal higher (90%)	No difference
v3-comp. math		No difference	No difference
v3-reas. math		No difference	No difference
		Older samples	
o2-comp. math		Verbal higher	Verbal higher
o2-reas. math		Verbal higher	Verbal higher
v2-comp. math		Verbal higher[a]	Space higher (90%)
v2-reas. math		No difference	Verbal higher (90%)
o3-comp. math		Verbal higher[a]	No difference
o3-reas. math		No difference	No difference
v3-comp. math		Verbal higher	Verbal higher
v3-reas. math.		Verbal higher	No difference

[a]A one-study result.

Tables 11 and 12 in the appendix give the statistics for overall comparisons of verbal-math and space-math correlations and of comparisons when correlations are divided by age of sample and math level, respectively.

Results for question two indicated that reasoning mathematics tasks are better related to spatial tasks than are computational tasks. A glance at numerical values suggests that this is true for verbal-math correlations as well: that is, reasoning math tasks have higher correlations with verbal tasks than have computational tasks. Table 4 presents numerical values of weighted averages for combined verbal-math correlations divided by math level: groups of verbal-math correlations have also been separated according to the spatial category of the correlations to which they were originally compared.

TABLE 4. Numerical values of weighted averages of combined verbal-math correlations, separated by level of mathematics task and spatial category.

		Mixed gender	Female	Male
o2	reasoning math	.59	.59	.58
	computational math	.42	.42	.40
v2	reasoning math	.48	.42[a]	.39
	computational math	.43	.43	.26
o3	reasoning math	.39	.41	.41
	computational math	.21	.39	.12
v3	reasoning math	.56	.56	.54
	computational math	.42	.44	.36

[a]This is the only pair of values in which the verbal-reasoning math correlation is lower than the verbal-computational math correlation.

<u>Question five</u>: <u>Are orientation spatial skills better related to mathematical problem-solving tasks than visualization spatial skills?</u> When mathematical tasks are classified as problem-solving according to test titles, visualization skills are as well-related to mathematical problem-solving tasks as are orientation skills. All numerical estimates suggest that visualization-math correlations are higher than orientation-math correlations. These results hold for samples regardless of gender

composition. Moreover, the results recur when orientation tasks are restricted to those in three dimensions: i.e., when only "pure" spatial skill is considered.

However, if the term "problem-solving" is stretched to included "mathematical reasoning," three-dimensional orientation tasks are well-related to SAT-Q tasks in the sense that they have the highest combined correlations -- e.g., .67 for basically corrected correlations in mixed-gender samples -- in the meta-analysis. However, the difference between these and the visualization-SAT-Q correlations is not statistically significant. (There are no studies which have two-dimensional orientation-SAT-Q correlations in the meta-analysis.) Tables 13 and 14 in The appendix contains the statistics for comparisons of orientation-problem solving and visualization-problem solving studies and for orientation-SAT-Q and visualization-SAT-Q studies, respectively, using mixed-gender samples and corrected correlations.

Question six: Do geometry, problem-solving, or SAT-Q mathematical tasks correlate more highly with spatial tasks than other mathematical tasks? For mixed gender samples, correlations were first divided into geometry-space and nongeometry-space correlations. Comparisons of these often found geometry-space correlations lower -- never higher -- than nongeometry-space correlations. Space-geometry proof correlations were then omitted from the geometry-space correlations, and the remaining correlations compared to nongeometry-space correlations. Comparisons then found differences insignificant except that internal comparisons showed geometry-v3 correlations higher than nongeometry-v3 correlations. Space-geometry proof correlations are significantly lower than other space-geometry correlations. Tables 15 and 16 in the appendix hold the statistics for these comparisons, using corrected correlations.

In single gender samples no significant differences are found except for a few external comparisons in which nongeometry-space correlations are higher than geometry-space correlations.

Tasks explicitly defined by researchers as mathematical problem-solving tasks do not produce math-space correlations which differ from other math-space correlations. In fact, numerical differences calculated are among the smallest of all those produced in this research. Table 17 in the appendix holds internal comparisons for mixed gender samples, using corrected correlations.

The College Board's Scholastic Aptitude Test, Quantitative Section (SAT-Q), produced unexpected results. In every spatial category represented, and for mixed gender, female, and male samples, external comparisons of corrected correlations indicate that space-SAT-Q correlations are higher than other space-math correlations, excepting only v3-SAT-Q comparison for male samples. (There were no o2-SAT-Q correlations in these data.) Internal comparisons produce no significant differences. However, the numerical differences are among the largest produced in this work; also, only four studies are available for internal comparisons: external comparisons may be the evidence of choice here. Table 5 displays external comparisons for mixed-gender samples.

Table 5. External comparisons of SAT-Q-space with other math-space z-transforms for three spatial categories.

	z^+	k	H	ψ	$\sigma^2(\xi)$	$\sigma^2(\psi)$	$\psi_1 - \psi_2$	CI
				o2-SAT-Q contrasted with other o2-math				
				No studies available.				
				v2-SAT-Q contrasted with other v2-math				
v2-SAT-Q	.620	6	7.8[b]	.620	0	.0017		
v2-math	.414	46	240.8	.429	.0252	.0008	.191	(.093,.289)
				o3-SAT-Q math contrasted with o3-math				
o3-SAT-Q	.782	3	11.6	.776	.0186	.0075		
o3-math	.397	37	176.9	.423	.0216	.0009	.353	(.174,.533)
				v3-SAT-Q math contrasted with v3-math				
v3-SAT-Q	.630	4	3.1[b]	.630	0	.0022		
v3-math	.508	52	282.8	.474	.0057	.0002	.156	(.060,.252)

NOTE. z^+ = weighted average of the z-transforms: k = number of studies; H = homogeneity statistic; ψ = estimate of the mean of the random effects model; $\sigma^2(\xi)$ = variance of the random effects model; $\sigma^2(\psi)$ = variance of the estimate of the mean of the random effects model; $\psi_1 - \psi_2$ = difference in means of the random effects model; CI = 95% confidence interval for difference in $\psi_1 - \psi_2$.

[b]Homogeneity upheld at the 95% confidence level.

It should be noted that in uncorrected correlations, the difference appears only in the three-dimensional orientation category. Because people who take the SAT-Q are always part of a selected sample, it can be argued that this is the population, and correlations shouldn't be corrected: however, if correlations are to be comparable, considering SAT-Q populations as subpopulations of the whole is reasonable.

Uncorrected correlations are not the statistics of choice to compare with corrected ones in SAT-Q data: all of these samples are fairly select, and it is reasonable that their correlations should require correction. Basically corrected correlations are a better comparison in checking that results from corrected correlations are not artefactual, even though the numbers of correlations in comparisons are smaller. The smaller number of studies probably increased the variance of random effects means, so not as many results were significant: however, enough of the results were significant to indicate that SAT-Q correlations are often higher than other math-space correlations. Table 6 gives a summary of the statistics using basically corrected correlations. Table 18 in the appendix holds the statistics themselves.

TABLE 6. Summary of results of external comparisons of SAT-Q-space and other math-space correlations, using basically corrected correlations

	Mixed gender	Females	Males
o2-math	No studies available		
v2-math	No difference	No difference	SAT-Q-space higher
o3-math	No difference	SAT-Q-space higher	SAT-Q-space higher
v3-math	SAT-Q-space higher	SAT-Q-space higher	No difference

Question seven: Gender difference calculations of correlations divided by spatial category produced some of the smallest means of the meta-analysis. None were significantly different from zero.

When groups are additionally subdivided by level of mathematical task and by age, some differences do appear: females sometimes have higher computational math-space correlations, and males sometimes have higher reasoning math-space correlations. Table 7 contains a summary of these gender difference results.

TABLE 7: Summary of gender difference calculations for correlations of spatial and various types of mathematics tasks.

	All mathematical tasks	
o2-math	No difference	
v2-math	No difference	
o3-math	No difference	
v3-math	No difference	
	Computational math	**Reasoning math**
	Samples of all ages	
o2-math	No difference	No difference
v2-math	No difference	No difference
o3-math	No difference	No difference
v3-math	Females higher	No difference
	Young samples	
o2-math	No difference	No difference
v2-math	No difference	Females higher, 90%
o3-math	No difference	Males higher
v3-math	Females higher	No difference
	Older samples	
o2-math	Females higher	No difference
v2-math	No difference	No difference
o3-math	No difference	No difference
v3-math	Females higher, 90%	Males higher, 90%

There are no gender differences in geometry-space or problem-solving-space correlations. However, females' SAT-Q correlations are higher than those of males in all samples, as well as in young and older samples. Uncorrected correlations - better comparison statistics in this case as it is reasonable to assume that female and male samples from the same studies will be similar with regard to selectivity , and thus all studies compared in corrected correlations can be used -- substantiated these results.

Tables 19-23 in the appendix contain the statistics for calculations of all gender differences noted above.

Conclusions

The general, theoretical questions posed in the introduction are addressed here.

Is there a particular spatial character to mathematical thought? The bulk of correlational evidence casts doubt on the conjecture that spatial skill is pervasive in mathematics as mathematics is taught and tested today. Mean correlations are low. Individual space-math correlations have been corrected for restriction of range, raising weighted averages and random effects model means. Yet the latter range from .35 to .47, numbers generally considered small or moderate (see, e.g., Cohen, 1987; DeVore, 1982).

The sequence of mean correlations suggests that correlations are higher for three-dimensional than two-dimensional spatial tasks and for visualization than orientation. Yet the highest of the average correlations is less than .50, and only the lowest and the highest are significantly different.

Can these low correlations be the result of uneven development of skills? Mathematical skills are trained; with the exception of two-dimensional orientation skills, spatial skills are not. However, by high school years, two-dimensional orientation and possibly two-dimensional visualization skills have ordinarily been developed: two-dimensional orientation skills are often part of the middle school curriculum. Three-dimensional skills may have received some emphasis in high school geometry. Yet correlations of older students are no higher than those of younger students except, marginally, in three-dimensional orientation. Differences calculated for correlations of different age groups are among the smallest in absolute value that we have seen here. From the sequence of mean correlations, it appears that the more familiar the spatial task, the smaller will be its correlation with mathematical tasks.

One of Smith's assertions found support in this meta-analysis: spatial tests do correlate more highly with reasoning than with computational mathematics tests, especially for males, and this differentiation appears more often in older students. However, even reasoning math-space correlations are not large. Moreover, informal comparisons indicate that verbal tests also correlate more highly with high-than with low-level mathematics tests. This last result is not consistent with the picture drawn by Smith.

The most damaging findings of this meta-analysis for the hypothesis that spatial ability underlies mathematical thought is that verbal-math correlations are often significantly higher -- almost never significantly lower -- than space-math correlations. Overall numerical estimates of mean differences between verbal-math and space-math correlations in each spatial category were always negative, indicating that verbal-math correlations were higher. The differences were significant for the two-dimensional spatial categories. They were significant for three-dimensional orientation in both groups of all-female and of all-male studies.

Considering male achievement at different ages, the results here contradict Smith's assertion that high spatial rather than high verbal ability is necessary for advanced mathematical courses. Verbal-math correlations were sometimes higher than space-math ones for older males, though almost never for younger ones. The picture was different for females. However, females seem to be irrelevant to Smith's argument: even before Sherman's conjecture was published, Smith wrote that females learn mathematics poorly precisely because of inadequate facility in imagery (see, e.g., 1964, p. 133).

Why was Smith so convinced that spatial and mathematical skills are outcomes of the same thought processes? He did not form his opinion from considering studies of college youth performing college mathematics, for there are few in the literature. Samples of college age youth often consist of education students or armed services personnel, and the mathematics tests used are not of college-level material. Hills' 1957 study is one that does concern college mathematics; Project Talent is another, though it reports on advanced mathematics as learned by high school students; Dick's study of college youth is another. The space-math correlations reported in Dick's study, corrected or not, are extremely low, not more than .25; correlations from the other two studies are not high.

Geometrical concepts are those which most immediately seem to involve spatial skill, and Smith reported some results which indicated that geometry-space correlations were higher than other mathematics-space correlations. The combined results of this meta-analysis do not agree with this conclusion. Geometry-space correlations were unexpectedly low, lower or at least no higher, than other math-space correlations.

A plausible explanation for this phenomenon is grounded in the fact that often, in geometry tasks, problems are presented to the student visually. Spatial skill may not be crucial when pictures are provided. Visual aids are helpful to the understanding of a problem: however, students should be able to, and can be taught to, generate visual aids for themselves if they are to succeed in solving problems presented to them without visuals (see Wong, 1988).

<u>Is there a substantial relationship between holistic spatial skill and mathematical problem solving skill?</u> If mathematical problem solving skill is measured by test items developed by researchers and test-constructors for that purpose -- items posed verbally which draw on computational, algebraic, geometric or combinatoric mathematical skills -- then the answer to the second question must again be negative. Mean combined correlations are not high: they are no higher for problem-solving tasks than for other mathematical tasks no matter whether orientation or visualization skills are considered. When orientation-math and visualization-math correlations are contrasted, the visualization-math correlations are higher numerically, though not, generally speaking, significantly so.

Shifting perspective slightly, we may take mathematical problem solving skill to include what is often referred to as "mathematical reasoning ability" -- the skill that SAT-Q administrators intend to test. Many SAT-Q items are multi-step, and thus could be expected to have more in common with visualization than with orientation tasks. However, there is some evidence that three-dimensional orientation is particularly strongly related to SAT-Q scores. The weighted average z-transform for the three o3-SAT studies was .782, the largest combined space-math z-transform in the meta-analysis. When external comparisons were made of orientation-SAT-Q and visualization-SAT-Q correlations, the numerical estimate of the difference favored orientation correlations, though, again, the difference was not significant.

The three o3-SAT studies all test older students. Gallagher's (1987) study was of gifted high school students: it produced the lowest correlations of the three. Johnson's (1984) study is of college students. He used the Guilford-Zimmerman "Clocks" test, which has been classified here as a three-dimensional orientation test, as all the items involve rotation of three-dimensional whole objects. However, more than one rotation is often involved, so the classification is somewhat problematic. The Burnett, Lane and Dratt study (1979) is of students at Rice University, a highly selective institution. The authors report that approximately 70% of those admitted will probably follow studies in engineering and science.

With these samples, intervening variables may be just as likely to explain the high correlation as some underlying connection of spatial and mathematical skills. The majority of students in the Burnett et al. study may have decided to concentrate on the sciences at a relatively early age, and thus become proficient in both mathematical and spatial skills. All subjects involved are certain to have had successful precollege experience, which should instill confidence in test-taking. Specialization and confidence should be explored as alternative explanations for the SAT-Q results.

This meta-analysis produced little evidence that holistic spatial skills are well-related to mathematical problem solving. Spatial strategies useful for mathematical problem solving often do not make heavy demands on spatial reasoning skill, though students confident with spatial reasoning may be more likely to attempt spatial strategies. More important than a high level of spatial ability is the conviction that these techniques have utility. The teacher's word that they do is not enough. Apparently students can be taught to generate visual aids in the process of mathematical problem solving. Educators should, perhaps, focus on this process in order to improve mathematical skills.

Does the relationship of spatial and mathematical skills differ for males and females? From the many calculations made, few significant or large gender differences appear in this study. Two dissimilarities between the genders in patterns of correlations emerge. The first concerns differentiation between correlations of spatial tasks with computational and reasoning mathematics tasks. Smith, Werdelin and others have suggested that spatial ability is involved in higher-level mathematical reasoning tasks as opposed to numerical, or computational tasks.

Females show less differentiation than males: Younger samples differentiate high and low-level mathematics tasks in two-dimensional visualization correlations, and older samples differentiate in both visualization categories. However, younger or older, they show no differentiation in orientation categories, and younger samples show no differentiation in either of the three-dimensional categories. There is some evidence that younger males do differentiate in three-dimensional categories; older males differentiate in every category.

This pattern difference is often the result of females' relatively high space-computational math correlations. It is consistent with Tartre's (1990) remark that females with low spatial ability have trouble with many tasks involved in problem solving: That spatial ability predicts computational skill just as reliably as reasoning mathematics skill may indicate that the two tend to be found together -- that is, that they are themselves well related in females.

The second pattern dissimilarity appears in the contrast of verbal-math and space-math correlations. When correlations are divided with regard to math level and age, females often have higher verbal-math correlations than space-math correlations whereas males tend to display this difference only in older samples. This could be construed as evidence that females are more verbal-analytic in approach than males. It could equally support the conclusion that school learning is more influential in mathematics learning for females than males.

Gender difference calculations provided only one significant difference -- females have higher SAT-Q-space correlations than males. This was true for both younger and older samples, using both corrected and uncorrected correlations.

An argument can be made that gender differences in favor of females could be expected in all math-space correlations when males have greater spatial skill than females and when spatial skill underlies abstract mathematical thought to a certain degree. In this scenario, spatial skill would be a better predictor of mathematical skill for females, because male spatial skill would be high for so many that small variations would not be reflected in mathematical achievement. (This is a version of the explanation of low correlations by uneven development of skills.) However, the gender difference in correlations in favor of females is unique to the SAT-Q correlations. And while generalizations can be made about the items of the SAT-Q, none of them transparently involve spatial skill.

Many of the items on the SAT-Q demand two or more mathematical observations for their solution. Confidence is almost surely a factor in the ability to carry out chains of reasoning: doubts can be debilitating at any link. Another generalization applies to the SAT-Q: it is a college entrance test. Students who take it have, on the average, performed well in high schools or college preparatory schools. High school students may have become specialized in their interests by the time they take the SAT: that is, they may have chosen to concentrate on the sciences or the humanities. Specialization seems particularly likely for those students applying to elite colleges. Common interest can also explain high correlations in the studies of gifted junior high school students who take the SAT, especially in the case of females.

Specialization and confidence have particular implications in the context of gender differences. Mathematical and spatial interests are each atypical for females. If atypical interests tend to cluster in the same individuals, it may be that mathematical and spatial interests are found together, developed to more or less the same degree, more often in females than males. This explanation is relevant to the younger SAT-Q-takers, who are classified as gifted students. It applies as well to the older females applying for college entry, whose atypical interests may have solidified into career ambitions. If spatial ability is developed to a substantial degree in any male, spatial rank is less likely to predict mathematical rank; in females, those who have opted for scientific careers will have developed spatial and mathematical abilities together, while those who have remained in the humanities may have been able to escape the development of either.

Concerning confidence, studies of gender differences in mathematics have found confidence in ability to be one of the most consistent predictors of achievement and participation in mathematics for both females and males. Such studies report consistent gender differences in favor of males with regard to confidence in mathematical ability (Benbow, 1988; Meyer and Koehler, 1990). This gender difference appears even in the Study of Mathematically Precocious Youth's seventh and eighth graders. The argument for confidence as an intervening variable producing the gender difference is similar to the one given for spatial ability: as males have more confidence, small variations would not be reflected in mathematical or spatial rankings as much for them as it would be for females. The argument is more convincing with regard to confidence than to spatial ability: standard deviations of measures of confidence are often smaller for males than for females, unlike standard deviations of cognitive measures (e.g., see Sherman and Fennema, 1977).

<u>Are differences in spatial skill the source of gender differences in mathematical tasks?</u> Overall, spatial skill does not appear to be a strong and pervasive influence in mathematical processes as they are taught and tested today. Moreover, of the areas of mathematics for which gender differences appear -- problem solving, geometrical tasks, and performance on mathematics sections of college entrance exams, only SAT-Q-space correlations provide any evidence of the involvement of spatial skill in mathematics.

We have suggested two alternative explanations for both the high correlations and the gender difference: one is confidence, leading to persistence in solving puzzles or handling chains of reasoning in abstract areas, and the other is common interests in non-verbal topics leading to a more parallel development of mathematical and spatial skills in females than males. The first of these alternatives takes account of a special character of the SAT-Q and is not impacted by the lack of gender differences found in other correlational categories as is the argument that spatial skill underlies abstract mathematical thought. The second alternative, common interest or specialization, is similarly unaffected by other findings in young age groups -- groups of subjects too young to have developed common interests -- and in studies of unselected samples.

In summary, correlational studies provide little evidence that spatial skill underlies abstract mathematical thought: correlations of spatial and mathematical tasks are moderate or small, certainly smaller than verbal-mathematical correlations. Predicting mathematical skill from spatial skill, or vice versa, is clearly an uncertain endeavor. Educators are not likely to be successful in improving performance on mathematical tasks as they are taught and tested today simply by improving spatial skills. Nor is the improvement of spatial skill likely to promote gender equity on mathematical tasks. There is evidence that we would be better off considering ways of training our students to generate nonverbal but simple representations of mathematical problems.

This is not to say that mathematics will always be taught and tested as it is today, nor that spatial skill is unimportant to gender equity in other sciences. However, alternative strategies must be found to correct the problems that exist in mathematical performance today.

NOTE

[1]Separate regressions were carried out for three types of samples: samples containing both female and male subjects (mixed-gender samples), samples of female subjects, and samples of male subjects. Early hopes were that the same equation would fit all three types, but this was not the case. Fisher z-transforms were used. The original z-transforms were always the strongest predictors. The mixed-gender and female samples produced similar equations. Where $\underline{S}$ denotes selectivity of sample, $\underline{N}$ denotes nationality of sample, $\underline{z}$ denotes the uncorrected z-transform, and $\underline{Z}$ denotes the corrected z-transform, for mixed-gender samples, the equation was

$$\begin{array}{llllllll} \underline{Z} = & -0.26 & + & (0.12)\underline{S} & + & (0.14)\underline{N} & + & (1.07)\underline{z}. \\ & (0.043) & & (0.016) & & (0.048) & & (0.062) \end{array}$$

The equation for female samples was

$$\begin{array}{llllll} \underline{Z} = & -0.19 & + & (0.10)\underline{S} & + & (1.01)\underline{z}. \\ & (0.060) & & (0.027) & & (0.136) \end{array}$$

The equation for males differed considerably. Where $\underline{A}$ denotes age and $\underline{Y}$ represents year of testing, the equation was

$$\begin{array}{llllllll} \underline{Z} = & -0.48 & + & (0.023)\underline{A} & + & (0.008)\underline{Y} & + & (1.08)\underline{z}. \\ & (0.141) & & (0.008) & & (0.002) & & (0.155) \end{array}$$

REFERENCES

Acker, M. B. (1968). The relation of achievement need, time perspective, and field articulation to academic performance (Doctoral dissertation, University of California at Berkeley, 1967). Dissertation Abstracts International, 29, 1492B. SUMA

Adams, V. M. and McLeod, D. B. (1979). The interaction of field dependence-independence and the level of guidance of mathematics instruction. Journal for Research in Mathematics Education, 10, 347-355. SUMA

Anglin, L., Meyer, R. and Wheeler, J. (1975). The relationship between spatial ability and mathematics achievement at the 4th and 6th grade levels (Working Paper No. 115). Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning. SUMA

Baldwin, S. L. (1985). Instruction in spatial skills and its effect on math achievement in the intermediate grades (Doctoral dissertation, University of Northern Colorado, 1984). Dissertation Abstracts International, 46, 595A. SUMA

Barakat, M. K. (1951). A factorial study of mathematical abilities. British Journal of Psychology (Statistical Section), 4, 137-156.

Battista, M. (1981). The interaction between two instructional treatments of algebraic structures and spatial-visualization ability. Journal for Educational Research, 74, 337-341.

Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. Journal for Research in Mathematics Education, 21, 47-60. SUMA

Battista, M. T., Wheatley, G. H., and Talsma, G. (1982). The importance of spatial visualization and cognitive development for geometry learning in preservice elementary teachers. Journal for Research in Mathematics Education, 13, 332-340. SUMA

Becker, B. J. (1978). The relationship of spatial ability to sex differences in the performance of mathematically precocious youths on the mathematical sections of the scholastic aptitude test. Unpublished master's thesis, Johns Hopkins University, Baltimore. SUMA

NOTE: References to studies used in the meta-analysis are followed by the acronym SUMA.

Becker, J. P. (1983). A study of the interactions of certain cognitive aptitudes with programmed figural and symbolic treatments of mathematical inequalities. Unpublished research report, Southern Illinois University at Carbondale, Carbondale, Illinois. SUMA

Behr, M. J. and Eastman, P. M. (1975). Interactions between structure-of-intellect factors and two methods of presenting concepts of modulus seven arithmetic -- a follow-up and refinement study. Journal for Research in Mathematics Education, 6, 150-157. SUMA

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, 169-232.

Bennett, G. K., Seashore, H. G. and Wesman, A. G. (1973). Differential Aptitude Tests. New York: Psychological Corporation.

Berry, P. C. (1958). An exploration of the interrelations among some non-intellectual predictors of achievement in problem solving (Tech. Rep. No. 4). New Haven: Yale University, Department of Industrial Administration and Department of Psychology. SUMA

Berry, P. C. (1959). A second exploration of the interrelations among some non-intellectual predictors of achievement in problem solving (Tech. Rep. No. 5). New Haven: Yale University, Department of Industrial Administration and Department of Psychology. SUMA

Bieri, J., Bradburn, W. M., and Galinsky, M. D. (1958). Sex differences in perceptual behavior. Journal of Personality, 26, 1-12. SUMA

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

Campos, A. and Cofan, E. (1986). Rotation of images and primary mental abilities: Influence of information and sex. Perceptual and Motor Skills, 63, 644-646. SUMA

Carment, D. G. (1988). The role of field independence/dependence, spatial visualization, and cerebral dominance in mathematics achievement of tenth grade students (Doctoral dissertation, Oklahoma State University, 1988). Dissertation Abstracts International, 49, 2957A. SUMA

Carry, L. R. (1968). Interaction of visualization and general reasoning abilities with instructional treatment in algebra (Doctoral dissertation, Stanford University, 1968). Dissertation Abstracts International, 29, 475A. SUMA

Cattell, R. B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton, Mifflin Company.

Childers, G. F. (1987). An error analysis of algebra problems containing exponents when constants, variables and algebraic expressions are interchanged in the laws, with implications for cognitive style theory and latent trait test-scoring methods (Doctoral dissertation, University of Virginia, 1987). Dissertation Abstracts International, SUMA

Cohen, J. (1987). Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Connor, J. M. and Serbin, L. A. (1985). Visual-spatial skill: Is it important for mathematics? Can it be taught? In S. F. Chipman, L. R. Brush, and D. M. Wilson (Eds.), Women and Mathematics: Balancing the Equation, (pp. 151-174). Hillsdale, NJ: Lawrence Erlbaum Asociates. SUMA

Cramond, B. L. (1982). Predicting mathematics achievement of gifted adolescent females (Doctoral dissertation, University of Georgia, 1982). Dissertation Abstracts International, 43, 2281A. SUMA

Davis, G. A. L. (1984). Hemispheric laterality, sex, visuospatial and verbal processing proficiency as predictors of mathematical ability in young children (Doctoral dissertation, University of Connecticut, 1983). Dissertation Abstracts International, 45, 106A. SUMA

DeVore, J. L. (1982). Probability and Statistics for Engineering and the Sciences. Monterey, California: Brooks/Cole Publishing Company.

DeWolfe, V. A. (1981). High school mathematics preparation and sex differences in quantitative abilities. Psychology of Women Quarterly, 5, 555-567. SUMA

Dick, T. (1984). Covariance structural models for mathematics achievement and participation: An investigation of sex differences at the level of college calculus using factorial modeling (Doctoral dissertation, University of New Hampshire, 1984). Dissertation Abstracts International, 45, 1673A. SUMA

Dubois, T. E. and Cohen, W. (1970). Relationship between measures of psychological differentiation and intellectual ability. Perceptual and Motor Skills, 31, 411-416. SUMA

Dunn, O. J. and Clark, V. (1969). Correlation coefficients measured on the same individuals. Journal of the American Statistical Association, 64, 366-377.

DuRapau, V. J., Jr. and Carry, L. R. (1981). Interaction of general reasoning ability and processing strategies in geometry instruction. Journal for Research in Mathematics Education, 12, 15-26. SUMA

Dziak, N. J. (1986). Programming computer graphics and the development of concepts in geometry (Doctoral dissertation, Ohio State University, 1985). Dissertation Abstracts International, 46, 2473B. SUMA

Ekstrom, R. B., French, J. W., Harmon, H. H., and Derwin, D. (1976). Kit of reference tests for cognitive factors. Princeton, New Jersey: Educational Testing Service.

Eastman, P. M and Carry, L. R. (1975). Interaction of spatial visualization and general reasoning abilities with instructional treatment in quadratic inequalities: a further investigation. Journal for Research in Mathematics Education, 6, 142-149.

Ellis, W. T. (1978). Piagetian development level and sex-role identification as factors in problem-solving performance and cognitive style (Doctoral dissertation, Southern Illinois University, 1977). Dissertation Abstracts International, 38, 6008A. SUMA

Elman, A. (1978). Within- and between-mode translation in Euclidean geometry (Doctoral dissertation, Stanford University, 1978). Dissertation Abstracts International, 39, 771A. SUMA

Emm, M. E. (1959). A factorial study of the problem-solving ability of fifth-grade boys (Doctoral dissertation, Catholic University of America, 1959). Washington, D. C.: The Catholic University of America Press. SUMA

Ethington, C. A. and Wolfle, L. M. (1984). Sex differences in a causal model of mathematics achievement. Journal for Research in Mathematics Education, 15, 361-377. SUMA

Farr, R. S. (1969). Personality variables and problem solving performance: An investigation of the relationships between field-dependence-independence, sex-role identification, problem difficulty and problem solving performance (Doctoral dissertation, New York University, 1968). Dissertation Abstracts International, 29, 2561A-2562A. SUMA

Fennema, E. (1980). Sex-related differences in mathematics achievement: Where and why? In L. H. Fox, L. Brody, and D. Tobin (Eds.), Women and the Mathematical Mystique (pp. 74-93). Baltimore: Johns Hopkins Press.

Fennema, E. and Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and affective factors. American Educational Research Journal, 14, 51-71. SUMA

Fennema, E. and Sherman, J. A. (1978). Sex-related differences in mathematics achievement and related factors: A further study. Journal for Research in Mathematics Education, 9, 189-203. SUMA

Fennema, E. and Tartre, L. A. (1985). The use of spatial visualization in mathematics by girls and boys. Journal for Research in Mathematics Education, 16, 184-206.

Filella, J. F. (1960). Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America. Genetic Psychology Monographs, 61, 116-163. SUMA

Flanagan, J. C., Davis, F. B., Dailey, J. T., Shaycoft, M. F., Orr, D. B., Goldberg, I., Neyman, C. A., Jr. (1964). The identification, development, and utilization of human talents: The American high-school student (Final report for Cooperative Research Project No. 635, U. S. Office of Education, Department of Health, Education, and Welfare). Pittsburgh: University of Pittsburgh, Project Talent Office. SUMA

French, J. W. (1957). The factorial invariance of pure-factor tests. Journal of Educational Psychology, 48, 93-109. SUMA

French, J. W. (1963). The relationship of problem-solving styles to the factor composition of tests (Research bulletin RB-63-15). Princeton: Educational Testing Service. SUMA

Friedlander, A. (1985). Achievement in similarity tasks: Effect of instruction, and relationship with achievement in spatial visualization at the middle grades level (Doctoral dissertation, Michigan State University, 1984). Dissertation Abstracts International, 45, 3570A-3571A. SUMA

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in quantitative tasks. Review of Educational Research, 59, 185-213.

Friedman, L. (1992). A Meta-Analysis of Correlations of Spatial and Mathematical Tasks. Unpublished doctoral dissertation, University of Chicago. (Available from the University.)

Gage, R. L. (1976). A study of the effects of positive and negative instances on the acquisition of selected algebra concepts as a function of cognitive style (Doctoral dissertation, University of Houston, 1976). Dissertation Abstracts International, 37, 4929A-4930A. SUMA

Gallagher, S. A. (1987). An analysis of visual-spatial ability, intellectual efficiency, and learning style on mathematics achievement of gifted male and gifted female adolescents (Doctoral dissertation, University of North Carolina at Chapel Hill, 1987). Dissertation Abstracts International, 48, 1733A. (University Microfilms No. 87-22294). SUMA

Galton, F. (1918). Inquiries into human faculty and its development (2nd ed.). New York: E. P. Dutton and Co.

Gardner, H. (1983). Frames of mind: the theory of multiple intelligences. New York: Basic Books.

Gardner, R. W., Jackson, D. N., and Messick, S. J. (1960). Personality organization in cognitive controls and intellectual abilities. Psychological Issues, 2, (Whole No. 8). SUMA

Gill, N. T., Herdtner, T. J., and Lough, L. (1968). Perceptual and socio-economic variables, instruction in body-orientation, and predicted academic success in young children. Perceptual and Motor Skills, 26, 1175-1184. SUMA

Glass, G. V., McGaw, B., and Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, California: Sage Publications.

Guay, R. B. and McDaniel, E. D. (1977). The relationship between mathematics achievement and spatial abilities among elementary school children. Journal for Research in Mathematics Education, 8, 211-215. SUMA

Guilford, J. P. (1965). Fundamental statistics in psychology and education. New York: McGraw-Hill Book Company.

Guilford, J. P., Comrey, A. L., Green, R. F. and Christensen, P. R. (1950, 1951). A factor-analytic study of reasoning abilities: I. Hypotheses and description of tests. II. Administration of tests and analysis of results (Reports Nos. 1 and 3). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Wilson, R. C., Christensen, P. R. and Lewis, D. J. (1951, 1952). A factor-analytic study of creative thinking: I. Hypotheses and description of tests. II. Administration of tests and analysis of results (Reports Nos. 4 and 8, Studies of Aptitudes of High-level Personnel). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Green, R. F., Christensen, P. R., Hertzka, A. F. and Kettner, N. W. (1952). A factor-analytic study of navy reasoning tests with the air force aircrew classification battery (Report No. 6, Studies of Aptitudes of High-level Personnel).

Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Hertzka, A. F., Berger, A. M. and Christensen, P. R. (1952, 1953). A factor-analytic study of evaluative abilities. I. Hypotheses and description of tests. II. Administration of tests and analysis of results (Reports Nos. 7, 9, Studies of Aptitudes of High-level Personnel). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Kettner, N. W. and Christensen, P. R. (1955). The relation of certain thinking factors to training criteria in the U.S. Coast Guard Academy (Report No. 13, Studies of Aptitudes of High-level Personnel). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Frick, J. W., Christensen, P. R., and Merrifield, P. R. (1957). A factor-analytic study of flexibility in thinking (Report No. 18, Studies of Aptitudes of High-level Personnel). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Guilford, J. P., Kettner, N. W. and Christensen, P. R. (1955). A factor-analytic investigation of the factor called general reasoning (Report No. 14, Studies of Aptitudes of High-level Personnel). Los Angeles: University of Southern California, Reports from the Psychological Laboratory. SUMA

Gulliver, M. P. (1979). The effect of the spatial visualization factor on achievement in operations with fractions (Doctoral dissertation, University of Minnesota, 1978). Dissertation Abstracts International, 39, 5381A-5382A. SUMA

Hanson, R. B. (1971). The prediction of achievement in high school plane geometry by spatial orientation-visualization-visual perception ability (Doctoral dissertation, University of Maryland, 1971). Dissertation Abstracts International, 32, 6003A. SUMA

Harris, M. G. and Harris, C. W. (1973). A structure of concept attainment abilities. Wisconsin Monograph Series. Madison Wisconsin: University of Wisconsin, Wisconsin Research and Development Center for Cognitive Learning. SUMA

Hedges, L. V. and Olkin, I. (1985). Statistical Methods for Meta-Analysis. Orlando: Academic Press, Inc.

Herot, L. W. (1976). The effects of a transformation approach to teaching geometry to low-achieving high school students (Doctoral dissertation, Columbia University Teachers' College, 1976). Dissertation Abstracts International, 37, 3478A-3479A. SUMA

Hesemann, J. P. (1976). A spatial model for the cognitive representation of verbal algebra problems (Doctoral dissertation, Indiana university, 1976). Dissertation Abstracts International, 37, 2037A. SUMA

Hills, J. R. (1957). Factor-analyzed abilities and success in college mathematics. Educational and Psychological Measurement, 17, 615-622. SUMA

Hilton, T. L. (1985). National changes in spatial-visual ability from 1960 to 1980 (Research Rep. # RR-85-27). Educational Testing Service, Princeton, New Jersey.

Hussien, G. A. (1979). Interactions between spatial and verbal abilities and two methods of presenting modulus seven arithmetic (Doctoral dissertation, Florida State University, 1979). Dissertation Abstracts International, 41, 144A. SUMA

Hyde, J. S., Fennema, E., and Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. Psychological Bulletin, 106,139-155.

Hyde, J. S., Geiringer, E. R., and Yen, W. M. (1975). On the empirical relation between spatial ability and sex differences in other aspects of cognitive performance. Multivariate Behavioral Research, 1975, 289-309. SUMA

Johnson, E. S. (1984). Sex differences in problem solving. Journal of Educational Psychology, 76, 1359-1371. SUMA

Jowett, B. (1937). The Dialogues of Plato, Vol. 1. New York: Random House.

Kelley, T. L. (1928). Crossroads in the mind of man. Stanford, California: Stanford University Press.

Kiser, O. L. (1986). Spatial ability and field mode in a computer-enhanced presentation of linear inequalities: an aptitude-treatment interaction study (Doctoral dissertation, University of Florida, 1986). Dissertation Abstracts International, 48, 860A. SUMA

Kline, W. E. (1959). A synthesis of two factor analyses of intermediate algebra. Psychometrika, 24, 343-359. SUMA

Koussy, A. A. H. El (1935). The visual perception of space. British Journal of Psychology, Monograph Supplement No. 20.

Krutetskii, V. A. (1976). The Psychology of Mathematics Abilities in Schoolchildren. Chicago: The University of Chicago Press.

Lean, G. and Clements, M. A. (1981). Spatial ability, visual imagery, and mathematical performance. Educational Studies in Mathematics, 12, 267-299. SUMA

Likert, R., and Quasha, W. H. (1970). Manual for the revised Minnesota Paper Form Board test. New York: The Psychological Corporation.

Linn, M. C. and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Development 56, 1479-1498.

Lunneborg, P. W. and Lunneborg, C. E. (1984). Contribution of sex-differentiated experiences to spatial and mechanical reasoning abilities. Perceptual and Motor Skills, 59, 107-113. SUMA

Lunneborg, P. W. and Lunneborg, C. E. (1986). Everyday spatial activities test for studying differential spatial experience and vocational behavior. Journal of Vocational Behavior, 28, 135-141. SUMA

Lynchard, B. B. (1988). The relationship of computational estimation to ability and selected variables of sixth grade students (Doctoral Dissertation, Delta State University, 1988). Dissertation Abstracts International, 49, 1686A. SUMA

Maccoby, E. E., and Jacklin, C. N. (1974). The Psychology of Sex Differences. Stanford: Stanford University Press.

MacGregor, S. K., Shapiro, J. Z., and Niemiec, R. (1988). Effects of a computer-augmented learning environment on math achievement for students with differing cognitive style. Journal of Educational Computing Research, 4, 453-465. SUMA

Manley, M. B. (1965). A factor analytic study of three types of concept attainment tasks (Doctoral dissertation, Princeton University). Dissertation Abstracts, XXVII, 600-B. SUMA

McCall, J. R. (1955). Sex differences in intelligence: a comparative factor study (Doctoral dissertation, Catholic University of America, 1955). Washington, D. C.: Catholic University of America. SUMA

McFie, J. (1961). The effect of education on African performance on a group of intellectual tests. British Journal of Educational Psychology, 31, 232-240. SUMA

McKay, L. R. (1979). Sex differences in cognitive styles and mathematics achievement in fourth and eighth graders (Doctoral dissertation, University of Texas at Austin, 1978). Dissertation Abstracts International, 39, 4103A. SUMA

McLeod, D. B. and Adams, V. M. (1979). The interaction of field independence with discovery learning in mathematics. Journal of Experimental Education, 48, 32-35. SUMA

McLeod, D. B. and Adams, V. M. (1980). Aptitude-treatment interaction in mathematics instruction using expository and discovery methods. Journal for Research in Mathematics Education, 11, 225-234. SUMA

McLeod, D. B. and Briggs, J. T. (1980). Interactions of field independence and general reasoning with inductive instruction in mathematics. Journal for Research in Mathematics Education, 11, 94-103. SUMA

McLeod, D. B., Carpenter, T. P., McCornack, R. L., and Skvarcius, R. (1978). Cognitive style and mathematics learning: The interaction of field independence and instructional treatment in numeration systems. Journal for Research in Mathematics Education, 9, 163-174. SUMA

McTaggart, H. P., M. A. (1959). A factorial study of the problem-solving ability of fifth-grade girls (Doctoral dissertation, Catholic University of America, 1959). Washington, D. C.: The Catholic University of America Press. SUMA

Meehan, A. M. (1984). Proportional reasoning as a function of sex, spatial ability, task spatial content, and task format (Doctoral dissertation, Temple University, 1984). Dissertation Abstracts International, 45, 377B-378B. SUMA

Meyer, R. A. (1978). Mathematical problem-solving performance and intellectual abilities of fourth-grade children. Journal for Research in Mathematics Education, 9, 334-348. SUMA

Meyer, M. R. and Koehler, M. S. (1990). Internal influences on gender differences in mathematics. In E. Fennema and G. Leder (Eds.), Mathematics and Gender (pp. 60-95). New York: Teachers College Press.

Michael, W. B., Guilford, J. P., Fruchter, B., and Zimmerman, W. S. (1957). The description of spatial-visualization abilities. Educational and Psychological Measurement, 17, 185-199.

Middaugh, D. J. (1980). Spatial ability and its relationship to the mathematical performance of adolescents (Doctoral dissertation, Kent State University, 1979). Dissertation Abstracts International, 40, 5349A. SUMA

Moore, T. L. and Witmer. J. A. (1991). Statistics within departments of mathematics at liberal arts colleges. American Mathematical Monthly, 98, 431-436.

Moses, B. E. (1977). The nature of spatial ability and its relationship to mathematical problem solving (Doctoral dissertation, Indiana University, 1977). Ann Arbor, Michigan: University Microfilms No. 77-30, 309. SUMA

Olson, J. F. (1985). Causes and correlates of mathematics anxiety and mathematics achievement: A path analytic approach (Doctoral dissertation, University of Nebraska -- Lincoln, 1985). Dissertation Abstracts International 46, 2975A. SUMA

Osburn, H. G. and Melton, R. S. (1963). Prediction of proficiency in a modern and traditional course in beginning algebra. Educational and Psychological Measurement, 23, 277-287. SUMA

Pattison, P. and Grieve, N. (1984). Do spatial skills contribute to sex differences in different types of mathematical problems? Journal of Educational Psychology, 76, 678-689. SUMA

Pearson, J. L. and Ferguson, L. R. (1989). Gender differences in patterns of spatial ability, environmental cognition, and math and English achievement in late adolescence. Adolescence, 24, 421-432. SUMA

Pedersen, K., Elmore, P., and Bleyer, D. (1986). Parent attitudes and student career interests in junior high school. Journal for Research in Mathematics Education, 17, 49-59. SUMA

Perunko, M. A. (1983). The relationships among mental imagery, spatial ability, analytic-synthetic processing and performance on mathematics problems (Doctoral dissertation, University of Maryland, 1982). Dissertation Abstracts International, 44, 1716A. SUMA

Phelps, E. and Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. Journal of Educational Psychology, 81, 639-646. SUMA

Porter, C. F. (1983). The effects of two-dimensional, three-dimensiona, and a combination of two- and three-dimensional instructional strategies oan the achievement of high school biology students with varying levels of spatial visualization ability, cognitive development, and gender (Doctoral dissertation, University of Georgia, 1982). Dissertation Abstracts International, 43, 2302A. SUMA

Punwar, A. (1970). Spatial visualization, reading, spelling, and mathematical abilities in second- and third-grade children. The American Journal of Occupational Therapy, 24, 495-499. SUMA

Ramirez, A. I. (1982). Modality and field dependence/independence: Learning style components and their relationship to mathematics achievement in the elementary school (Doctoral dissertation, Florida State University, 1982). Dissertation Abstracts International, 43, 666A. SUMA

Roberge, J. J. and Flexer, B. K. (1981). Re-examination of the covariation of field independence, intelligence and achievement. British Journal of Educational Psychology, 51, 235-236. SUMA

Rogers, A. L. (1918). Experimental tests of mathematical ability and their prognostic value. New York: Teachers' College (Contributions to Education, No. 89).

Ronning, R. R., McCurdy, D. and Ballinger, R. (1984). Individual differences: a third component in problem-solving instruction. Journal of Research in Science Teaching, 21, 71-82. SUMA

Rosenberg, J. H. (1981). The ability of selected cognitive, affective, and educational variables to predict the presence of anxiety related to mathematics (Doctoral dissertation, University of Connecticut, 1981). Dissertation Abstracts International, 41, 3890A. SUMA

Rosenfeld, I. J. (1958). Mathematical ability as a function of perceptual field-dependency and certain personality variables (Doctoral dissertation, University of Oklahoma, 1958). Dissertation Abstracts, 19, 880-881. SUMA

Ruszel, Rev. H. A. (1952). Test patterns in intelligence: Comparative factor analyses for high school boys and girls (Doctoral dissertation, Catholic University of America, 1952). Washington, D. C.: The Catholic University of America Educational Research Monographs. SUMA

Satterly, D. J. (1976). Cognitive styles, spatial ability, and school achievement. Journal of Educational Psychology, 68, 36-42. SUMA

Schoenfeld, A. (1985). Mathematical Problem Solving. Orlando: Academic Press, Inc.

Schonberger, A. K. (1976). The interrelationship of sex, visual spatial abilities, and mathematical problem solving ability in grade seven (Doctoral dissertation, University of Wisconsin -- Madison, 1976). Ann Arbor: University Microfilms No. 76-20, 132. SUMA

Semmes, P. A. (1978). A preliminary investigation of two presentation modes on dynamic problems in algebra in relation to individual differences (Doctoral dissertation, University of Texas at Austin, 1977). Dissertation Abstracts International, 38, 4005A-4006A. SUMA

Shaycoft, M. F., Dailey, J. T., Orr, D. B., Neyman, C. A. Jr. and Sherman, S. E. (1963). The identification, development, and utilization of human talents: Studies of a complete age group -- Age 15 (Report of the Cooperative Research Project No. 566, U. S. Office of Education, Department of Health, Education, and Welfare). Pittsburgh: University of Pittsburgh, Project Talent Office. SUMA

Sherman, J. A. (1967). The problem of sex differences in space perception and aspects of intellectual functioning. Psychological Review, 74, 290-299.

Sherman, J. and Fennema, E. (1977). The study of mathematics by high school girls and boys: Related variables. American Educational Research Journal, 14, 159-168.

Shieh, W. (1986). Spatial visuzlization, attitudes toward mathematics, and mathematics achievement among Chinese-American, Hispanic-American, and Caucasian seventh and eighth grade students (Doctoral dissertation, University of the Pacific, 1985). Dissertation Abstracts International, 46, 3633A. SUMA

Shipman, V. C. (1971). Disadvantaged children and their first school experiences: Structure and development of cognitive competencies and styles prior to school entry (Research bulletin No. PR-71-19, Educational Testing Service, Head Start Longitudinal Study). Princeton, NJ: Educational Testing Service. (ERIC Document No. 068 522.) Information also received from Technical Reports Nos. 4 and 6, Technical Report series, Educational Testing Service, Head Start Longitudinal Study (ERIC Document No. 081 813). SUMA

Smith, I. M. (1964). Spatial ability: Its education and social significance. London: University of London Press Ltd.

Solan, H. A. (1987). The effects of visual-spatial and verbal skills on written and mental arithmetic. Journal of the American Optometric Association, 58, 88-94. SUMA

Spearman, C. (1927). The abilities of man. London: MacMilian.

Stake, R. E. (1961). Learning parameters, aptitudes, and achievements (Pschometric Monograph No. 9). Richmond, Virginia: The William Byrd Press, Inc. SUMA

Sweeney, E. J. (1953). Sex differences in problem solving (Technical Report No. 1, prepared under contract N6onr 25125 for the Office of Naval Research) (NR 153-149). Stanford, California: Stanford University, Department of Psychology. SUMA

Tabler, M. B. (1981). The relationship between one dimension of cognitive style and students' mathematics performance (Doctoral dissertation, Indiana University, 1980). Dissertation Abstracts International, 41, 4281A. SUMA

Tartre, L. A. (1990). Spatial skills, gender, and mathematics. In E. Fennema and G. Leder (Eds.), Mathematics and Gender (pp. 27-59). New York: Teachers College Press.

Thurstone, L. L. (1938). Primary mental abilities. Chicago: University of Chicago Press.

Tillotson, M. L. (1985). The effect of instruction in spatial visualization on spatial abilities and mathematical problem solving (Doctoral dissertation, University of Florida, 1984). Dissertation Abstracts International, 45, 2792A. SUMA

Vaidya, S. and Chansky, N. (1980). Cognitive development and cognitive style as factors in mathematics achievement. Journal of Educational Psychology, 72, 326-330. SUMA

Vandenberg, Steven G. (1971). The Mental Rotation Test. Boulder, Colorado: Institute for Behavioral Genetics.

Very, P. S. (1967). Differential factor structures in mathematical ability. Genetic Psychology Monographs, 75, 169-207. SUMA

Webb, L. F. and Carry, L. R. (1975). Interaction of spatial visualization and general resoning abilities with instructional treatment in quadratic inequalities: A follow-up study. Journal for Research in Mathematics Education, 6, 132-140. SUMA

Weiner, N. C. (1984). Cognitive aptitudes, personality variables, and gender difference effects on mathematical achievement for mathematically gifted students (Doctoral dissertation, Arizona State University, 1983). Dissertation Abstracts International, 44, 3621A. SUMA

Werdelin, I. (1958). The mathematical ability: experimental and factorial studies (Studia Psychologica et Paedagogica, Series altera, Investigationes IX). Lund, Sweden: C. W. K. Gleerup. SUMA

Werdelin, I. (1961). Geometrical ability and the space factors in boys and girls (Studia Psychologica et Paedagogica [Lund Studies in Psychology and Education], Series altera, Investigationes X). Lund, Sweden: C. W. K. Gleerup. SUMA

Wilson, J. H. (1933). Group factors among abilities involved in a school certificate examination. British Journal of Educational Psychology, 3, 71-108.

Witkin, H. A., and Goodenough, D. R. (1981). Cognitive styles: Essence and origins: Field dependence and field independence. New York, International Universities Press, Inc.

Witkin, H. A., Moore, C. A., Oltman, P. K., Goodenough, D. R., Friedman, F., Owen, D. R. and Raskin, E. (1977). Role of the field-dependent and field-independent cognitive styles in academic evolution: A longitudinal study. Journal of Educational Psychology, 69, 197-211 (1977). SUMA

Wolfe, L. R. (1970). The effects of spatial visualization on spatial ability and arithmetic achievement in junior high school students (Doctoral dissertation, State University of New York at Albany, 1970). Dissertation Abstracts International, 31, 2801A. SUMA

Wong, S. K. P. (1988). The effects of different visual strategies, selected individual differences, and learner control on student word-problem solving performance (Doctoral dissertation, University of Minnesota, 1988). Dissertation Abstracts International, 49, 1366A. SUMA

Wormack, L. (1980). Sex differences in factorial dimension of verbal, logical, mathematical and visuospatial ability. Perceptual and Motor Skills, 50, 445-446. The following reference gives information about the same study. SUMA

Wormack, L. (1979). Cognitive predictors of articulation in writing. Perceptual and Motor Skills, 48, 1151-1156.

Wrigley, J. (1958). The factorial nature of ability in elementary mathematics. British Journal of Educational Psychology, 28, 61-78.

Young, C. D., Jr. and Becker, J. P. (1979). The interaction of cognitive aptitudes with sequences of figural and symbolic treatments of mathematical inequalities. Journal for Research in Mathematics Education, 10, 24-37. SUMA

Zimowski, M. F., and Wothke, W. (1986). The Measurement of Human Variation in Spatial Visualizing Ability: A Process-Oriented Perspective. Technical Report 1986-1. Chicago: Johnson O'Connor Research Foundation, Incorporated.

APPENDIX

	Page
Table 8	54
Table 9	55
Table 10	58
Table 11	62
Table 12	65
Table 13	71
Table 14	72
Table 15	73
Table 16	75
Table 17	76
Table 18	77
Table 19	79
Table 20	80
Table 21	82
Table 22	82
Table 23	83

TABLE 8. Random effects model means for space-math correlations divided by spatial category and by gender, for uncorrected and basically corrected correlations

	uncorrected				basically corrected			
	r^r	k	$\sigma^2(\psi)$	CI	r^r	k	$\sigma^2(\psi)$	CI
				Mixed-gender				
o2-math	.321	25	.0001	(.299,.342)	.350	8	.0001	(.332,.369)
v2-math	.375	52	.0004	(.341,.407)	.442	16	.0012	(.386,.496)
o3-math	.358	40	.0009	(.305,.409)	.568	7	.0093	(.427,.682)
v3-math	.396	55	.0002	(.370,.422)	.474	17	.0002	(.454,.492)
				Females				
o2-math	.298	16	.0004	(.262,.334)	.338	4	.0001	(.319,.356)
v2-math	.387	27	.0007	(.343,.493)	.397	7	.0118	(.204,.560)
o3-math	.311	23	.0034	(.244,.376)	.559	4	.0191	(.345,.717)
v3-math	.400	30	.0005	(.361,.438)	.507	12	.0008	(.465,.546)
				Males				
o2-math	.312	22	.0003	(.282,.341)	.331	5	.0010	(.274,.385)
v2-math	.340	36	.0003	(.307,.371)	.409	7	.0145	(.195,.585)
o3-math	.313	27	.0009	(.258,.313)	.543	5	.0145	(.357,.688)
v3-math	.371	42	.0003	(.341,.400)	.465	10	.0005	(.430,.498)

NOTE: See Table 1 for explanation of notation.

TABLE 9. Internal comparisons of computational and reasoning mathematics-space z-transforms, divided by spatial category, using corrected correlations

	z_{+1}	z_{+2}	z_{+dif}	k	H	ψ	CI
					Mixed-gender		
o2-math							
c = -1	.276	.368	-.093	12	12.7[b]	-.093	(-.118,-.067)
c = 0					25.4	-.095	(-.136,-.054)
c = 0.9					254.3	-.088	(-.124,-.053)
v2-math							
c = -1	.372	.485	-.113	17	18.1[b]	-.113	(-.177,-.048)
c = 0					36.1	-.104	(-.181,-.027)
c = 0.9					361.8	-.089	(-.162,-.017)
o3-math							
c = -1	.248	.477	-.230	4	12	-.199	(-.431,.032)
						90% CI	(-.394,-.005)
c = 0					24.1	-.193	(-.423,.037)
c = 0.9					240.9	-.187	(-.415,.041)
v3-math							
c = -1	.286	.506	-.220	16	85.9	-.160	(-.236,-.185)
c = 0					171.7	-.154	(-.226,-.083)
c = 0.9					1717.4	-.144	(-.208,-.080)

(Table continues)

				Females			
o2-math							
c = -1	.328	.347	-.019	6	9.2[a]	-.019	(-.057,.018)
c = 0					18.4	-.032	(-.098,.034)
c = .9					184.3	-.042	(-.102,.018)
v2-math							
c = -1	.283	.425	-.142	8	5.3[a]	-.142	(-.256,-.027)
c = 0					10.7[a]	-.142	(-.222,-.061)
c = .9					106.8	-.169	(-.274,-.064)
o3-math							
c = -1	.194	.380	-.186	4	6.5[a]	-.186	(-.353,-.019)
c = 0					12.9	-.178	(-.433,.078)
c = .9					129.3	-.168	(-.420,.085)
v3-math							
c = -1	.326	.487	-.161	11	29.5	-.115	(-.190,-.040)
c = 0					59.1	-.106	(-.179,-.033)
c = .9					590.8	-.098	(-.163,-.033)
				Males			
o2-math							
c = -1	.236	.334	-.098	13	10.1[a]	-.098	(-.134,-.063)
c = 0					20.1[b]	-.124	(-.167,-.080)
c = .9					201.1	-.152	(-.190,-.113)
v2-math							
c = -1	.226	.333	-.106	15	16.3[a]	-.106	(-.176,-.037)
c = 0					32.6	-.101	(-.183,-.019)
c = .9					326.3	-.091	(-.170,-.013)

(Table continues)

Males (cont.)

o3-math							
c = -1	.150	.342	-.192	6	5.7[a]	-.192	(-.300,-.084)
c = 0					11.5	-.180	(-.301,-.059)
c = .9					114.8	-.172	(-.291,-.052)
v3-math							
c = -1	.231	.474	-.242	20	45.4	-.256	(-.318,-.194)
c = 0					90.8	-.268	(-.328,-.207)
c = .9					908.3	-.281	(-.335,-.228)

NOTE: $\underline{z}^{+1}$ = average weighted mean of z-transforms from the first group; $\underline{z}^{+2}$= average weighted mean of z-transforms from the second group;
$\underline{z}^{+dif}$ = average weighted mean of the difference of z-transforms from the first and second groups; $\underline{k}$ = number of studies; $\underline{H}$ = homogeneity statistic for $\underline{z}^{+dif}$ in this group; ψ = random effects model mean for the difference in z-transforms; CI = 95% confidence interval for ψ.

[a]The numbers in columns 1-4 are the same for all values of c. This is not obvious in the case of $\underline{z}^{+dif}$: however, as

$$\underline{z}^{+dif} = \sum[[(\underline{n}_i-3)/2(1-\underline{c})]/[\sum(\underline{n}_j-3)/2(1-\underline{c})]]\,(\underline{z}_{i1} - \underline{z}_{i2}),$$

the constant factor 2(1-$\underline{c}$) cancels; thus $\underline{c}$ plays no role in the calculations.

[b]Homogeneity upheld at the 95% confidence level.

Table 10. Internal comparisons of space-math z-transforms for different spatial categories between z-transforms of computational and reasoning mathematics tests, by gender, divided by age of sample, using corrected correlations.

	z^{+1}	z^{+2}	z^{+dif}	k	H	ψ	CI
					Young samples		
					o2-math		
Females							
c = -1	.375	.329	.046	3	0.9[a]	.046	(-.138,.231)
c = 0					1.8[a]	.046	(-.084,.177)
c = .9					17.5	.090	(-.041,.222)
Males							
c = -1	.447	.512	-.065	3	0.06[a]	-.065	(-.267,.136)
c = 0					0.1[a]	-.065	(-.208,.077)
c = .9					1.1[a]	-.065	(-.110,-.020)[c]
					v2-math		
Females							
c = -1	.347	.437	-.090	6	1.4[a]	-.090	(-.215,.035)
c = 0					2.9[a]	-.090	(-.179,-.002)
c = .9					28.6	-.094	(-.167,-.022)
Males							
c = -1	.327	.364	-.036	6	0.3[a]	-.036	(-.168,.096)
c = 0					0.5[a]	-.036	(-.130,.057)
c = .9					5.1[a]	-.036	(-.066,-.007)[d]

(Table continues)

				o3-math			
Females							
c = -1	.262	.305	-.043	2	0.04[a]	-.043	(-.293,.208)
c = 0					0.09[a]	-.043	(-.219,.135)
c = .9					0.9[a]	-.043	(-.097,.014)
Males							
c = -1	.300	.421	-.120	2	0.03[a]	-.120	(-.391,.150)
c = 0					0.06[a]	-.120	(-.311,.068)
c = .9					0.6[a]	-.120	(-.181,-.060)[e]
				v3-math			
Females							
c = -1	.556	.572	-.016	4	3.8[a]	-.016	(-.131,.098)
c = 0					7.7[b]	-.063	(-.214,.089)
c = .9					76.6	-.084	(-.232,.064)
Males							
c = -1	.468	.545	-.077	4	2.9[a]	-.077	(-.198,.044)
c = 0					5.8[a]	-.077	(-.163,.009)
					90% CI	-.077	(-.149,-.005)
c = .9					57.7	-.129	(-.268,.009)
					90% CI		(-.246,-.012)

(Table continues)

Older samples

				o2-math			
Females							
c = -1	.329	.349	-.021	5	8.3[b]	-.029	(-.092,.035)
c = 0					16.5	-.039	(-.105,.027)
c = .9					165.4	-.076	(-.137,-.015)[f]
Males							
c = -1	.229	.328	-.099	10	9.9[a]	-.099	(-.136,-.063)
c = 0					19.8	-.135	(-.085,-.086)
c = .9					197.8	-.175	(-.218,-.130)
				v2-math			
Females							
c = -1	-.028	.367	-.395	2	0.06[a]	-.395	(-.673,-.117)
c = 0					0.1[a]	-.395	(-.592,-.199)
c = .9					1.2[a]	-.395	(-.457,-.333)
Males							
c = -1	.187	.320	-.133	9	14.6[b]	-.140	(-.261,-.019)
c = 0					29.1	-.138	(-.259,-.017)
c = .9					291.2	-.139	(-.255,-.022)
				o3-math			
Females							
c = -1	.139	.440	-.300	2	4.2	-.307	(-.763,.149)
c = 0					8.3	-.308	(-.764,.148)
c = .9					83.0	-.309	(-.765,.147)
Males							
c = -1	.121	.327	-.206	4	5.4[b]	-.200	(-.361,-.039)
c = 0					10.8	-.195	(-.356,-.034)
c = .9					107.8	-.190	(-.350,-.029)

(Table continues)

				v3-math			
Females							
c = -1	.304	.479	-.174	7	19.0	-.141	(-.219,-.062)
c = 0					38.0	-.129	(-.207,-.052)
c = .9					380.4	-.109	(-.179,-.038)
Males							
c = -1	.214	.469	-.255	16	34.9	-.280	(-.343,-.215)
c = 0					69.7	-.296	(-.358,-.234)
c = .9					697.4	-.318	(-.372,-.262)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

[b]Homogeneity disputed at the 90%, but not 95%, confidence level.

[c]95% confidence interval does not cover zero for $c \geq .75$.

[d]95% confidence interval does not cover zero for $c > .80$.

[e]90% confidence interval does not cover zero for $c \geq .45$.

[f]95% confidence interval does not cover zero for $c > .80$.

TABLE 11. Comparison of space-math and verbal math z-transforms in the four spatial categories

	$\underline{z}_{+1}$	$\underline{z}_{+2}$	$\underline{z}_{+dif}$	$\underline{k}$	$\underline{H}$	ψ	CI	
				Mixed-gender samples				
				o2-math contrasted with verbal-math				
c = -1	.351	.668	-.316	16	94.9	-.201	(-.283,-.119)	
c = 0					189.8	-.188	(-.265,-.110)	
c = 0.9					1898.1	-.168	(-.239,-.098)	
				v2-math contrasted with verbal-math				
c = -1	.388	.524	-.135	23	28.9[a]	-.135	(-.186,-.085)	
c = 0					58.0	-.131	(-.193,-.068)	c = 0.9
			579.6	-.122	(-.182,-.062)			
				o3-math contrasted with verbal-math				
c = -1	.350	.421	-.071	15	35.0	-.054	(-.153,.046)	
c = 0					70.0	-.046	(-.145,.053)	
c = 0.9					700.0	-.036	(-.133,.062)	
				v3-math contrasted with verbal-math				
c = -1	.469	.623	-.154	29	95.4	-.056	(-.115,.002)	
						90% CI	(-.105,-.007)	
c = 0					190.9	-.043	(-.098,.011)	
c = 0.9					1908.6	-.029	(-.077,.019)	

(Table continues)

Females

			o2-math contrasted with verbal-math				
c = -1	.330	.666	-.336	14	35.5	-.257	(-.341,-.174)
c = 0					70.9	-.230	(-.309,-.151)
c = .9					709.3	-.198	(-.266,-.130)
			v2-math contrasted with verbal-math				
c = -1	.390	.472	-.082	21	29.8[a]	-.082	(-.151,-.012)
c = 0					59.6	-.111	(-.205,-.016)
c = .9					596.7	-.120	(-.209,-.031)
			o3-math contrasted with verbal-math				
c = -1	.309	.462	-.153	15	21.0[a]	-.153	(-.243,-.063)
c = 0					42.1	-.126	(-.240,-.012)
c = .9					420.5	-.110	(-.222,.001)
						90% CI	(-.204,-.016)
			v3-math contrasted with verbal-math				
c = -1	.458	.627	-.169	23	48.7	-.115	(-.183,-.047)
c = 0					97.3	-.097	(-.162,-.032)
c = .9					973.2	-.063	(-.118,-.007)

(Table continues)

Males

			o2-math contrasted with verbal-math					
c = -1	.348	.651	-.303	24	124.4	-.168	(-.243,-.194)	
c = 0					248.8	-.156	(-.226,-.086)	
c = .9					2488.0	-.141	(-.204,-.078)	
			v2-math contrasted with verbal-math					
c = -1	.343	.443	-.099	39	58.6	-.107	(-.159,-.054)	
c = 0					117.2	-.111	(-.164,-.059)	c = .9
			1171.6	-.118	(-.168,-.068)			
			o3-math contrasted with verbal-math					
c = -1	.308	.388	-.079	22	38.4	-.074	(-.148,-.0003)	
c = 0					76.8	-.071	(-.145,.003)	
						90% CI	(-.133,-.009)	
c = .9					768.5	-.069	(-.141,.003)	
						90% CI	(-.130,-.008)	
			v3-math contrasted with verbal-math					
c = -1	.461	.593	-.132	43	143.5	-.028	(-.081,.024)	
c = 0					287.0	-.022	(-.071,.028)	
c = .9					2870.2	-.022	(-.066,.022)	

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

TABLE 12. Comparison of space-math and verbal-math z-transforms for young (< 14 years) and older (≥ 14 years) samples, divided by level of mathematics test and the four spatial types of spatial task, by gender.

	z^{+1}	z^{+2}	z^{+dif}	k	H	ψ	CI
o2-math contrasted with verbal-math, young samples, low math							
Females							
c = -1	.277	.611	-.334	2	2.6[a]	-.334	(-.584,-.083)
c = 0					5.1	-.213	(-.707,.282)
c = .9					51.3	-.187	(-.684,.311)
Males							
c = -1	.495	.416	.079	2	.25[a]	.079	(-.192,.349)
c = 0					.51[a]	.079	(-.113,.270)
c = .9					5.1	.111	(-.046,.267)
o2-math contrasted with verbal-math, older samples, low math							
Females:							
c = -1	.321	.457	-.136	3	2.2[a]	-.136	(-.174,-.098)
c = 0					4.4[a]	-.136	(-.163,-.109)
c = .9					44.5	-.133	(-.174,-.093)
Males:							
c = -1	.258	.435	-.177	6	7.7[a]	-.177	(-.215,-.140)
c = 0					15.4	-.162	(-.220,-.104)
c = .9					154.4	-.160	(-.214,-.107)

(Table continues)

o2-math contrasted with verbal-math, young samples, high math

<u>Females</u>

c = -1	.321	.635	-.313	4	10.5	-.306	(-.612,.001)
					90% CI		(-.564,-.048)
c = 0					20.9	-.303	(-.608,.002)
					90% CI		(-.567,-.039)
c = .9					209.0	-.300	(-.601,.002)
					90% CI		(-.554,-.046)

<u>Males</u>

c = -1	.436	.532	-.096	4	5.6[a]	-.096	(-.264,.072)
c = 0					11.2	-.068	(-.306,.171)
c = .9					112.3	-.050	(-.285,.186)

o2-math contrasted with verbal-math, older samples, high math

<u>Females</u>

c = -1	.331	.669	-.338	10	24.0	-.260	(-.345,-.176)
c = 0					48.1	-.220	(-.302,-.139)
c = .9					480.8	-.164	(-.234,-.095)

<u>Males</u>

c = -1	.330	.667	-.337	16	68.3	-.173	(-.267,-.078)
c = 0					136.6	-.152	(-.241,-.063)
c = .9					1366.0	-.132	(-.214,-.050)

v2-math contrasted with verbal-math, young samples, low math

<u>Females:</u>

c = -1	.295	.521	-.226	4	2.7[a]	-.226	(-.385,-.066)
c = 0					5.5[a]	-.226	(-.339,-.113)
c = .9					54.6	-.212	(-.381,-.042)

<u>Males:</u>

c = -1	.347	.490	-.143	4	9.8	-.226	(-.573,.121)
c = 0					19.6	-.254	(-.598,.090)
c = .9					196.4	-.284	(-.621,.054)

(Table continues)

v2-math contrasted with verbal-math, older samples, low math

Females							
c = -1	.424	.709	-.285	1	0	-.285	(-.501,.020)
					90% CI		(-.504,-.066)
c = 0					0	-.285	(-.501,-.069)
c = .9					0	-.285	(-.525,-.045)
Males							
c = -1	.290	.199	.091	4	1.3[a]	.091	(-.054,.236)
c = 0					2.6[a]	.091	(-.012,.193)
					90% CI		(.005,.177)
c = .9					25.5	.095	(-.009,.200)
					90% CI		(.007,.183)

v2-math contrasted with verbal-math, young samples, high math

Females:							
c = -1	.379	.592	-.212	9	6.7[a]	-.212	(-.325,-.099)
c = 0					13.4[b]	-.213	(-.326,-.100)
c = .9					134.1	-.224	(-.334,-.114)
Males:							
c = -1	.410	.555	-.146	10	4.8[a]	-.146	(-.254,-.038)
c = 0					9.5[a]	-.146	(-.222,-.069)
c = .9					95.2	-.133	(-.216,-.050)

v2-math contrasted with verbal-math, older samples, high math

Females								
c = -1	.321	.306	.017	10	10.9[a]	.017	(-.077,.111)	c = 0
			21.9	-.006	(-.127,.115)			
c = .9					218.6	-.009	(-.123,.106)	
Males								
c = -1	.285	.330	-.045	19	16.6[a]	-.045	(-.105,.016)	
c = 0					33.2	-.061	(-.125,.003)	
					90% CI		(-.115,-.007)	
c = .9					332.3	-.083	(-.144,-.022)	

(Table continues)

o3-math contrasted with verbal-math, young samples, low math

Females							
c = -1	.250	.611	-.361	2	2.3[a]	-.361	(-.612,-.110)
c = 0					4.6	-.251	(-.717,.215)
c = .9					45.7	-.223	(-.693,.247)
Males							
c = -1	.234	.416	-.182	2	0.4[a]	-.182	(-.452,.089)
c = 0					0.7[a]	-.182	(-.373,.009)
					90% CI		(-.343,-.021)
c = .9					6.7	-.142	(-.325,.042)

o3-math contrasted with verbal-math, older samples, low math

Females							
c = -1	.390	.710	-.320	1	0	-.320	(-.626,-.015)
c = 0					0	-.320	(-.536,-.105)
c = .9					0	-.320	(-.389,-.252)
Males							
c = -1	.213	.132	.080	2	1.3[a]	.080	(-.092,.252)
c = 0					2.6[a]	.080	(-.042,.202)
c = .9					26.1	.013	(-.227,.252)

o3-math contrasted with verbal-math, young samples, high math

Females							
c = -1	.312	.655	-.343	3	5.8[b]	-.308	(-.656,.041)
					90% CI		(-.601,-.015)
c = 0					11.6	-.295	(-.643,.053)
					90% CI		(-.588,-.002)
c = .9					115.9	-.281	(-.625,.063)
Males							
c = -1	.388	.438	-.049	4	1.8[a]	-.049	(-.211,.113)
c = 0					3.6[a]	-.049	(-.164,.065)
c = .9					36.2	-.032	(-.161,.097)

(Table continues)

o3-math contrasted with verbal-math, older samples, high math

Females							
c = -1	.252	.327	-.076	12	10.7[a]	-.076	(-.181,.030)
c = 0					21.4	-.053	(-.161,.054)
c = .9					213.9	-.031	(-.137,.075)
Males							
c = -1	.300	.363	-.062	16	29.0	-.067	(-.174,.041)
c = 0					58.0	-.068	(-.175,.039)
c = .9					580.4	-.066	(-.170,.039)

v3-math contrasted with verbal-math, young samples, low math

Females							
c = -1	.541	.587	-.046	4	14.4	-.109	(-.403,.186)
c = 0					28.7	-.104	(-.395,.187)
c = .9					287.2	-.094	(-.378,.191)
Males							
c = -1	.444	.498	-.054	4	3.8[a]	-.054	(-.176,.067)
c = 0					7.5[b]	-.078	(-.238,.083)
c = .9					75.0	-.065	(-.222,.092)

v3-math contrasted with verbal-math, older samples, low math

Females							
c = -1	.275	.453	-.177	6	6.0[a]	-.177	(-.215,-.140)
c = 0					11.9	-.162	(-.213,-.112)
c = .9					119.4	-.134	(-.182,-.086)
Males							
c = -1	.230	.417	-.186	10	20.0	-.141	(-.213,-.069)
c = 0					40.1	-.117	(-.187,-.046)
c = .9					401.0	-.093	(-.154,-.031)

(Table continues)

v3-math contrasted with verbal-math, young samples, high math

Females							
c = -1	.502	.527	-.025	7	20.6	-.028	(-.251,.196)
c = 0					41.2	-.017	(-.237,.202)
c = .9					411.8	-.008	(-.219,.203)
Males							
c = -1	.439	.453	-.014	8	5.4[a]	-.014	(-.112,.084)
c = 0					10.8[a]	-.014	(-.084,.055)
c = .9					108.0	-.025	(-.117,.068)

v3-math contrasted with verbal-math, older samples, high math

Females							
c = -1	.453	.640	-.187	14	19.3[a]	-.187	(-.222,-.152)
c = 0					38.7	-.138	(-.198,-.077)
c = .9					386.8	-.088	(-.139,-.037)
Males							
c = -1	.447	.613	-.166	27	99.5	-.040	(-.117,.038)
c = 0					198.9	-.031	(-.104,.043)
c = .9					1989.2	-.032	(-.099,.035)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

TABLE 13. Comparison of orientation-problem-solving z-transforms with visualization-problem solving transforms, for mixed gender samples, using corrected correlations.

	z^{+1}	z^{+2}	z^{+dif}	k	H	ψ	CI
Orientation-problem solving contrasted with visualization-problem solving							
	Or.	**Vis.**					
c = -1	.391	.455	-.064	14	15.8[a]	-.064	(-.137,.010)
					90% CI		(-.124,-.002)
c = 0					31.9	-.062	(-.151,.026)
c = 0.9					317.3	-.077	(-.162,.008)
					90% CI		(-.149,-.005)[b]
o3 -problem solving contrasted with visualization-problem solving							
c = -1	.388	.456	-.068	12	14.2[a]	-.068	(-.148,.011)
					90% CI		(-.135,-.001)
c = 0					28.4	-.059	(-.155,.037)
c = 0.9					284.1	-.046	(-.139,.046)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level

[b]Confidence interval does not cover zero for $c \geq .9$.

Table 14. External comparison of z-transforms of orientation and of visualization correlations with SAT-Q studies for mixed-gender samples, using corrected correlations

	z^{+}	k	H	ψ	$\sigma^2(\xi)$	$\sigma^2(\psi)$	$\psi_1 - \psi_2$	CI
Orien.	.782	3	11.6	.776	.0186	.0075		
Vis.	.629	8	10.3[a]	.629	0	.0011		
Difference in means							.147	(-.035,.329)

NOTE: See Table 5 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

TABLE 15. Internal comparisons of space-geometry z-transforms with other math-space z-transforms, using corrected correlations from mixed-gender samples

	$\underline{z}^{+1}$	$\underline{z}^{+2}$	$\underline{z}^{+dif}$	$\underline{k}$	$\underline{H}$	ψ	CI
			o2-geometry contrasted with other o2-math				
c = -1	.377	.489	-.112	5	1.8[a]	-.112	(-.217,-.006)
c = 0					3.6[a]	-.112	(-.186,-.037)
c = 0.9					35.9	-.124	(-.197,-.051)
			o2-geometry contrasted with other o2-math (no proof-space correlations included)				
c = -1	.456	.516	-.060	4	1.8[a]		(-.174,.054)
c = 0					3.6[a]		(-.141,.020)
c = 0.9					35.8		(-.166,.016)
			v2-geometry contrasted with other v2-math				
c = -1	.379	.521	-.143	4	0.3[a]		(-.281,-.004)
c = 0					0.6[a]		(-.240,-.045)
c = 0.9					5.8[a]		(-.174,-.112)
			v2-geometry contrasted with other v2-math (no proof-space correlations included)				
c = -1	.568	.576	-.008	3	1.4[a]		(-.185,.121)
c = 0					2.8[a]		(-.119,.104)
c = 0.9					27.3		(-.233,.099)
			o3-geometry contrasted with other o3-math				
c = -1	.384	.434	-.050	4	0.8[a]		(-.156,.056)
c = 0					1.6[a]		(-.125,.025)
c = 0.9					15.8	-.039	(-.088,.010)

(Table continues)

			o3-geometry contrasted with other o3-math (no proof-space correlations included)				
c = -1	.481	.449	.032	4	0.3[a]		(-.081,.146)
c = 0					0.5[a]		(-.048,.113)
c = 0.9					5.2[a]		(.007,.058)[b]
			v3-geometry contrasted with other v3-math				
c = -1	.393	.463	-.070	5	0.6[a]		(-.175,-.036)
c = 0					1.1[a]		(-.145,.005)
						90%CI	(-.133,-.007)
c = 0.9					10.7	-.064	(-.104,-.023)
			v3-geometry contrasted with other v3-math (no proof-space correlations included)				
c = -1	.589	.511	.078	4	0.9[a]		(-.036,.192)
c = 0					1.9[a]		(-.002,.159)
						90%CI	(.010,.146)
c = 0.9					18.7	.068	(.002,.134)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

[b]Intervals excluded zero only for c ≥ .85.

TABLE 16. Internal comparisons of geometry proof and nonproof z-transforms using corrected correlations from mixed-gender samples

	z^{+1}	z^{+2}	z^{+dif}	k	H	ψ	CI
			geometry proof-space contrasted with other geometry-space				
c = -1	.447	.614	-.167	4	0.3[a]	-.167	(-.286,-.048)
c = 0					0.5[a]	-.167	(-.251,-.083)
c = 0.9					5.4[a]	-.167	(-.194,-.140)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% level.

TABLE 17. Internal comparisons of space-problem solving with other space-math z-transforms, for mixed-gender samples, using corrected correlations.

	z+1	z+2	z+dif	k	H	Ψ	CI
				o2-problem solving contrasted with other o2-math			
c = -1	.438	.404	.033	4	2.1[a]	.033	(-.051,.118)
c = 0					4.3[a]	.033	(-.026,.093)
c = 0.9					43	.036	(-.020,.092)
				v2-problem solving contrasted with other v2-math			
c = -1	.499	.481	.018	13	4.9[a]	.018	(-.052,.088)
c = 0					9.8[a]	.018	(-.032,.067)
c = 0.9					97.7	.016	(-.032,.063)
				o3-problem solving contrasted with other o3-math			
c = -1	.437	.436	.002	7	3[a]	.002	(-.086,.089)
c = 0					6[a]	.002	(-.060,.064)
c = 0.9					60.2	-.022	(-.086,.041)
				v3-problem solving contrasted with other v3-math			
c = -1	.493	.477	.016	16	23.8[b]	.005	(-.063,.074)
c = 0					47.7	-.004	(-.072,.065)
c = 0.9					476.9	-.011	(-.075,.053)

NOTE: See Table 9 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

[b]Homogeneity disputed at the 90% confidence level, not at the 95% level.

TABLE 18. External comparisons of SAT-Q math-space with other math-space z-transforms for the four spatial categories, by gender, using basically corrected correlations

	z+	k	H	ψ	σ2(ξ)	σ2 (ψ)	ψ1 - ψ2	CI
o2 -SAT-Q contrasted with other o2-math								
No studies available.								
v2 -SAT-Q contrasted with other v2-math								
Females								
v2-SAT-Q	.541	4	19.4	.542	.1071	.0327		
v2-math	.244	3	0.2[a]	.244	0	.0027		
difference in means							.298	(-.071,.667)
Males								
v2-SAT-Q	.614	4	14.5	.564	.1074	.0340		
v2-math	.218	3	0.9[a]	.218	0	.0021		
difference in means							.346	(-.026,.718)
						90% CI		(.033,.660)
o3 -SAT-Q contrasted with other o3-math								
Females								
o3-SAT-Q	.719	3	13.3	.726	.0519	.0204		
o3-math	.302	1	0[c]	.302	0	.0200		
difference in means							.424	(.030,.818)
Males								
o3-SAT-Q	.761	3	16.0	.744	.0499	.0190		
o3-math	.392	2	0.001[a]	.392	0	.0044		
difference in means							.352	(.052,.652)

(Table continues)

v3 -SAT-Q contrasted with other v3-math

Females								
v3-SAT-Q	.910	3	3.1[a]	.910	0	.0065		
v3-math	.499	9	16.6	.512	.0010	.0003		
difference in means							.398	(.236,.560)
Males								
v3-SAT-Q	.545	3	6.0	.555	.0337	.0171		
v3-math	.488	7	17.2	.497	.0014	.0004		
difference in means							.058	(-.202,.317)

NOTE: See Table 5 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

TABLE 19. Gender difference calculations for space-math z-transforms in the four spatial categories, using corrected correlations

	$\underline{z}^+$	Mdif	$\underline{k}$	Hdif	ψ(dif)	CI
o2-math						
Females	.348	.0007	15	45.7	-.042	(-.106,.023)
Males	.345					
v2-math						
Females	.461	.057	24	71.0	.045	(-.051,.141)
Males	.390					
o3-math						
Females	.367	-.047	21	37.7	-.049	(-.125,.026)
Males	.415					
v3-math						
Females	.482	-.013	29	71.0	-.025	(-.074,.025)
Males	.484					

NOTE: $\underline{z}^+$ = weighted average mean for the group of z-transforms; Mdif = weighted average mean difference of female and male z-transforms from the same study; $\underline{k}$ = number of studies; Hdif = homogeneity statistic for Mdif; ψ(dif) = random effects model mean for the difference of female and male z-transforms from the same study; CI = 95% confidence interval for ψ(dif).

TABLE 20. Gender difference calculations for space-math z-transforms in the four spatial categories and for mathematics tasks separated by level, using corrected correlations

	z^{+}	Mdif	k	Hdif	ψ(dif)	CI	
				o2-computational math			
Females	.330	.067	9	20.6	.038	(-.020,.097)	**Males** .256
				o2-reasoning math			
Females	.348	.001	14	44.5	-.038	(-.103,.026)	**Males** .345
				v2-computational math			
Females	.330	.002	11	17.0[b]	-.016	(-.118,.086)	**Males** .270
				v2-reasoning math			
Females	.455	.060	21	68.1	.045	(-.060,.151)	**Males** .382
				o3-computational math			
Females	.290	-.016	6	5.2[a]	-.016	(-.115,.082)	**Males** .221
				o3-reasoning math			
Females	.353	-.054	19	33.9	-.061	(-.139,.017)	**Males** .413

(Table continues)

			v3-computational math				
Females	.331	.053	14	29.3	.058	(.010,.105)	**Males**
.264							
			v3-reasoning math				
Females	.481	-.013	25	70.9	-.027	(-.080,.026)	**Males**
.489							

NOTE: See Table 19 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

[b]Homogeneity disputed at the 90% confidence level, not at the 95% level.

Table 21. Gender difference calculations for space-geometry z-transforms using corrected correlations.

	z^+	Mdif	k	Hdif	ψ(dif)	CI
			Geometry nonproof			
Females	.439	.059	6	5.7[a]	.059	(-.053,.171)
Males	.355					
			Geometry proof			
Females	.278	.016	4	19.6	-.035	(-.337,.268)
Males	.252					

NOTE: See Table 19 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

TABLE 22. Gender difference calculations for space-problem solving z-transforms, using corrected correlations

	z^+	Mdif	k	Hdif	ψ(dif)	CI
Problem solving						
Females	.458	-.011	18	40.5	.027	(-.053,.107)
Males	.455					

NOTE: See Table 19 for explanation of notation.

TABLE 23. Gender difference calculations for SAT-Q--space z-transforms using corrected correlations

	z+	Mdif	k	Hdif	ψ(dif)	CI
			SAT-Q-space			
Females	.591	.145	12	18.6[b]	.137	(.021,.253)
Males	.457					
			SAT-Q-space, young samples			
Females	.690	.346	3	1.7[a]	.346	(.126,.567)
Males	.340					
			SAT-Q,-space, older samples			
Females	.582	.122	9	13.3[a]	.122	(.046,.198)
Males	.475					

NOTE: See Table 19 for explanation of notation.

[a]Homogeneity upheld at the 95% confidence level.

[b]Homogeneity disputed at the 90% confidence level, not at the 95% level.