Spatial Ability for STEM Domains: Aligning Over 50 Years of Cumulative Psychological Knowledge Solidifies Its Importance

Jonathan Wai, David Lubinski, and Camilla P. Benbow
Vanderbilt University

The importance of spatial ability in educational pursuits and the world of work was examined, with particular attention devoted to STEM (science, technology, engineering, and mathematics) domains. Participants were drawn from a stratified random sample of U.S. high schools (Grades 9–12, $N = 400,000$) and were tracked for 11+ years; their longitudinal findings were aligned with pre-1957 findings and with contemporary data from the Graduate Record Examination and the Study of Mathematically Precocious Youth. For decades, spatial ability assessed during adolescence has surfaced as a salient psychological attribute among those adolescents who subsequently go on to achieve advanced educational credentials and occupations in STEM. Results solidify the generalization that spatial ability plays a critical role in developing expertise in STEM and suggest, among other things, that including spatial ability in modern talent searches would identify many adolescents with potential for STEM who are currently being missed.

Keywords: spatial ability, talent searches, longitudinal study, STEM, constructive replication

Over 50 years ago, Super and Bachrach (1957) published *Scientific Careers,* a report of a National Science Foundation (NSF) advisory panel. Appearing the year *Sputnik* was launched, this document characterized the personal attributes of scientists and engineers for the purposes of better identifying human capital and, ultimately, uncovering ways to nurture scientific and technical potential. It also was the year of two landmark publications in the *American Psychologist*: Cronbach's (1957) APA Presidential Address, on "The Two Disciplines of Scientific Psychology," wherein the importance of tailoring educational interventions and opportunities to individual differences among students was emphasized, and Paterson's (1957) Bingham Lecture, "The Conservation of Human Talent," which reinforced this idea.

Emphasized throughout Super and Bachrach (1957) was the critical role of spatial ability, a construct aptly defined by Lohman (1994a, p. 1000) as "the ability to generate, retain, retrieve, and transform well-structured visual images." Spatial ability was characterized as an individual differences attribute with particular relevance for learning the advanced scientific–technical material needed for developing outstanding STEM (science, technology,

engineering, and mathematics) contributors, those individuals capable of moving engineering and physical science disciplines forward. However, in their review Super and Bachrach stressed that attributes beyond spatial ability—mathematical ability in particular, as well as interests and nonintellectual determinants such as persistence—should be studied also. They further voiced that "longitudinal studies beginning at a relatively early age and extending over a period of some 10 to 15 years seemed called for" (Super & Bachrach, 1957, p. 87). This study sequences two such longitudinal studies: one from 1960 to 1974 and a second that began in 1971 and is still ongoing.

Contemporary Neglect of Utilizing Psychological Knowledge About Spatial Ability

Part of the motivation for this article is that currently, over 50 years after Super and Bachrach's (1957) report, relatively little implementation of spatial ability is found for selection, curriculum, and instruction in educational settings—even in STEM domains, where it appears to be highly relevant. This neglect is especially surprising as we live in a globally competitive world (Friedman, 2005), and the need to identify and nurture scientific and technical talent has never been greater (American Competitiveness Initiative, 2006; National Academy of Sciences, 2005). Indeed, with plenty of evidence for the educational–occupational significance of spatial ability accumulated (Gohm, Humphreys, & Yao, 1998; Humphreys, Lubinski, & Yao, 1993; Lohman, 1988, 1994a, 1994b; Smith, 1964), Richard E. Snow (1999) expressed perplexity about the neglect of spatial ability in applied educational circles:

> There is good evidence that [spatial ability] relates to specialized achievements in fields such as architecture, dentistry, engineering, and medicine . . . . Given this plus the longstanding anecdotal evidence on the role of visualization in scientific discovery, . . . it is incredible that

there has been so little programmatic research on admissions testing in this domain. (p. 136)

Since Snow's (1999) observation, at least two promising studies have appeared that further underscore the importance of assessing spatial ability among intellectually talented youths initially identified by mathematical and verbal measures. These studies also suggest a venue wherein assessing spatial ability could have an immediate impact, because both were based on talent search participants (Benbow & Stanley, 1996; Colangelo, Assouline, & Gross, 2004; Stanley, 2000). (Talent search participants are young adolescents who take college entrance exams 4 years earlier than is typical in order to qualify for special educational programs for talented youths.) Talent searches could relatively easily add spatial ability measures to their selection criteria and thereby cast a wider net for identifying intellectually able youths for educational experiences in architecture, engineering, robotics, and the physical sciences. However, the assessment of spatial ability may benefit more students than just talented youths. Basic science indicates that students throughout the ability range could profit from spatial ability assessments and the provision of educational opportunities aimed at developing spatial ability (Humphreys et al., 1993; Humphreys & Lubinski, 1996; Lohman, 2005; Smith, 1964).

The two studies on spatial ability discussed above that appeared after Snow (1999) were based on independent cohorts of participants in the Study of Mathematically Precocious Youth (SMPY; Lubinski & Benbow, 2006). SMPY is a longitudinal study currently in its fourth decade and consisting of five cohorts identified at different time points. It is designed to uncover the best methods for identifying and nurturing talent for STEM as recommended by Super and Bachrach (1957). Shea, Lubinski, and Benbow (2001) tracked 563 talent search participants identified with the Scholastic Assessment Test (SAT) by age 13 as intellectually talented (top 0.5% for their age-group); at the time of their identification in the late 1970s, they were assessed on spatial ability also. Over a 20-year interval, biographical, educational, and occupational criteria were collected 5, 10, and 20 years after initial identification. Relative to criterion groupings in the humanities and other disciplines, the young adolescents who subsequently found math–science to be their favorite high school course, earned undergraduate and graduate degrees in STEM, and ultimately ended up in a STEM career 20 years later, typically displayed higher levels of spatial ability at age 13. Moreover, the discriminant function analyses conducted at all three time points revealed that spatial ability added *incremental validity* (accounted for a statistically significant amount of additional variance) beyond SAT-Mathematical (measuring mathematical reasoning ability) and SAT-Verbal (measuring verbal reasoning ability) in predicting these math–science criteria.

Subsequently, Webb, Lubinski, and Benbow (2007), using a less select talent search sample of 1,060 adolescents identified in the mid-1990s (top 3% in ability), provided evidence that spatial ability possesses incremental validity over both SAT scales and comprehensive educational–occupational preference questionnaires over a 5-year interval for predicting favorite high school course, leisure activities relevant to STEM, college major, and intended occupation (i.e., the predictive period spanned from initial identification at age 13 to after high school). Overall, spatial ability accounted for an additional 3% of the variance in predicting these criteria beyond both SAT measures and two comprehensive educational–vocational preference questionnaires. Again, relative to young adolescents whose outcomes fell in the humanities and in other disciplines, participants with STEM outcomes displayed higher levels of spatial ability at age 13.

As suggestive as these findings are, however, D. F. Lohman (personal communication, May 2007) noted one limitation: Shea et al. (2001) and Webb et al. (2007) were not based on random samples of the general population or even random samples of high-ability students. All participants in both studies were talent search participants, students identified as highly able who often were motivated to attend academically challenging programs for talented youths. Would spatial ability play a similar role among students not identified in this fashion? One purpose in our study is to provide an answer to this question. In addition, to solidify the length of time that spatial ability has been known to play a consistent role in the development of STEM expertise, we decided to try to bridge the gap between the studies reviewed in Super and Bachrach's (1957) NSF report and our contemporary findings from talent search participants (Shea et al., 2001; Webb et al., 2007). For this purpose, we explored Project TALENT, a massive longitudinal study launched just following Super and Bachrach's report (in 1960) and culminating with an 11-year follow-up in the early 1970s (Wise, McLaughlin, & Steel, 1979), when the first SMPY participants were identified. Figure 1 illustrates the bridge we are aiming to build.

Project TALENT consists of four cohorts totaling 400,000 participants. They were identified as high school students (Grades 9 through 12, approximately 100,000 per grade) shortly after Sputnik was launched. Subsequently, they were followed up 11 years after their high school graduation in the early 1970s, when modern talent searches for intellectually precocious youths were just being launched (Keating & Stanley, 1972; Stanley, 1996). Therefore, if the findings uncovered by this study of Project TALENT participants correspond with those of studies conducted prior to and reviewed in Super and Bachrach's (1957) NSF report, and, in addition, if they mirror modern findings based on talent search participants identified throughout the 1970s and 1990s and fol-

Fifty Years of Longitudinal Research on Spatial Ability

Figure 1. Over 50 years of cumulative empirical research on the educational and occupational significance of spatial ability for STEM.

lowed up in current times, the collective findings would establish a solid foundation for educational practice. This foundation would be derived from two distinctive longitudinal studies of the type Super and Bachrach (1957) called for (and which were launched over successive time frames): Project TALENT (1960 to early 1970s) and SMPY (early 1970s to present times).

Studies that have been conducted periodically for over 50 years with different populations and that consistently reveal similar patterns are rare in educational psychology. Furthermore, there is a methodological rationale for the importance of sequencing such studies. Following Lykken's (1968, 1991) nomenclature for conducting replications in psychological research, if all of these longitudinal studies mirror one another, aligning their findings over multiple decades would constitute a series of *constructive replications*, which are the most scientifically compelling kind: The idea behind constructive replication is to vary systematically as many construct-irrelevant design features as possible over successive replications, while ensuring that the focal construct is preserved in each study. In Lykken's (1968) words,

> To obtain an ideal constructive replication, one would provide a competent investigator with *nothing more than* a clear statement of the empirical "fact" which the first author would claim to have established . . . and then let the replicator formulate his own methods of sampling, measurement, and data analysis . . . . We are interested in the *construct*, . . . not the *datum*. (p. 156)

In the current context, the studies we are aligning employed different measures, cohorts, time points, longitudinal intervals, investigators, and criteria. Yet, the focal construct, *spatial ability*, and its role in various educational and occupational pursuits remained the same. With this foundation, the following study was conducted.

Logic and Constructive Replication Sample

The specific objectives in this study were (a) to determine the extent to which spatial ability has operated consistently for decades in the prediction of educational and occupational criteria with particular emphasis on STEM domains, (b) to determine the extent to which early manifestations of exceptional spatial ability portend the development of STEM expertise, and (c) to demonstrate how neglect of this important dimension of cognitive functioning leads to untapped pools of talent for STEM domains.

The Shea et al. (2001) findings constitute, to our knowledge, the first demonstration that spatial ability adds incremental validity (beyond mathematical and verbal ability measures) in the prediction of educational–occupational criteria among talent search participants initially identified before age 13 on the basis of SAT-Math and SAT-Verbal scores. Some of their longitudinal outcomes, which include favorite and least favorite high school course (age 18 follow-up), college major (age 23 follow-up), and occupation (age 33 follow-up), are shown in Figure 2, as a function of their standing on these three abilities assessed at age 13 in standard deviation units. Mathematical ability is scaled on the *x*-axis, verbal ability on the *y*-axis, and spatial ability on the *z*-axis (notated by arrows in standard deviation units; arrows to the right are positive effect sizes for spatial ability, and arrows to the left are negative effect sizes for spatial ability). Essentially, this is a three-dimensional graph put in a two-dimensional representation.

This figure will serve as a template for replication purposes. To visualize the location of each group in three-dimensional space, imagine the arrows to the right projecting outward (toward you) and the arrows to the left projecting inward (away from you), both perpendicular to the *x*- and *y*-axes; in this way, the psychological distance between these criterion groups can be pictured in the space defined by the three ability dimensions. Dotted lines are placed around the STEM groups to highlight their consistent pattern across all three time points. We predicted that these patterns also would be observed in Project TALENT participants, whose 11-year longitudinal follow-up was conducted before these SMPY participants were identified in the 1970s at age 13.

It is important to keep in mind that although the SMPY participants were identified as intellectually talented in early adolescence (top 0.5% for their age-group), their patterns of specific abilities are readily distinguished as a function of contrasting educational–occupational group membership. With respect to spatial ability, the focal construct under analysis here, the consistently above-average spatial ability of participants in STEM educational degree groupings and occupations reveals the importance of spatial ability in STEM arenas (as indicated by rightward-pointing arrows across all four panels of Figure 2). Within the dotted boxes in each panel of Figure 2 are the STEM groups. However, to clarify the graph, examine just the physical science group in Panel C. This group has a positive *z*-score value (relative to the other groups) on mathematical, verbal, and spatial ability. Stated differently, those individuals who majored in physical science had higher mathematical, verbal, and spatial abilities relative to those who majored in other areas. In contrast, examine the humanities group in Panel C. This group has a positive *z*-score value on verbal ability but a negative *z*-score value on both mathematical and spatial ability relative to the other groups. What this means is that those individuals who majored in the humanities had relatively higher verbal ability but relatively lower math and spatial ability in comparison to those with other majors.

Consistently lower levels of spatial ability, indicated by arrows pointing to the left, are associated with domains outside of STEM. For example, referring back to the physical science group in Panel C, this group's *z* score on spatial ability was 0.34 (the length of the rightward-pointing arrow), whereas the humanities group's *z* score on spatial ability was −0.34 (the length of the leftward-pointing arrow). This means that these two groups are 0.68 standard deviations apart on spatial ability, even though both groups were above the normative mean on spatial ability (Shea et al., 2001). Hence, relative strengths and weaknesses manifested during adolescence are related to contrasting outcomes in education and the world of work. Jointly, these successive panels demonstrate how spatial ability operates over the life span (after high school, after college, and at age 33), regardless of whether it is measured. That is, whereas mathematical and verbal ability measures were used to identify these participants and similar measures were subsequently used throughout their educational careers as selection tools, spatial ability was assessed experimentally only at the time of their initial identification; spatial ability was not then used and is very rarely currently used in educational selection for advanced degrees or professional careers. Spatial ability played a clear role for these intellectually talented youths in domains in which it is placed at a premium (as well as those in which it is not). Multiple examples of how spatial ability operated in the attainment of educational and
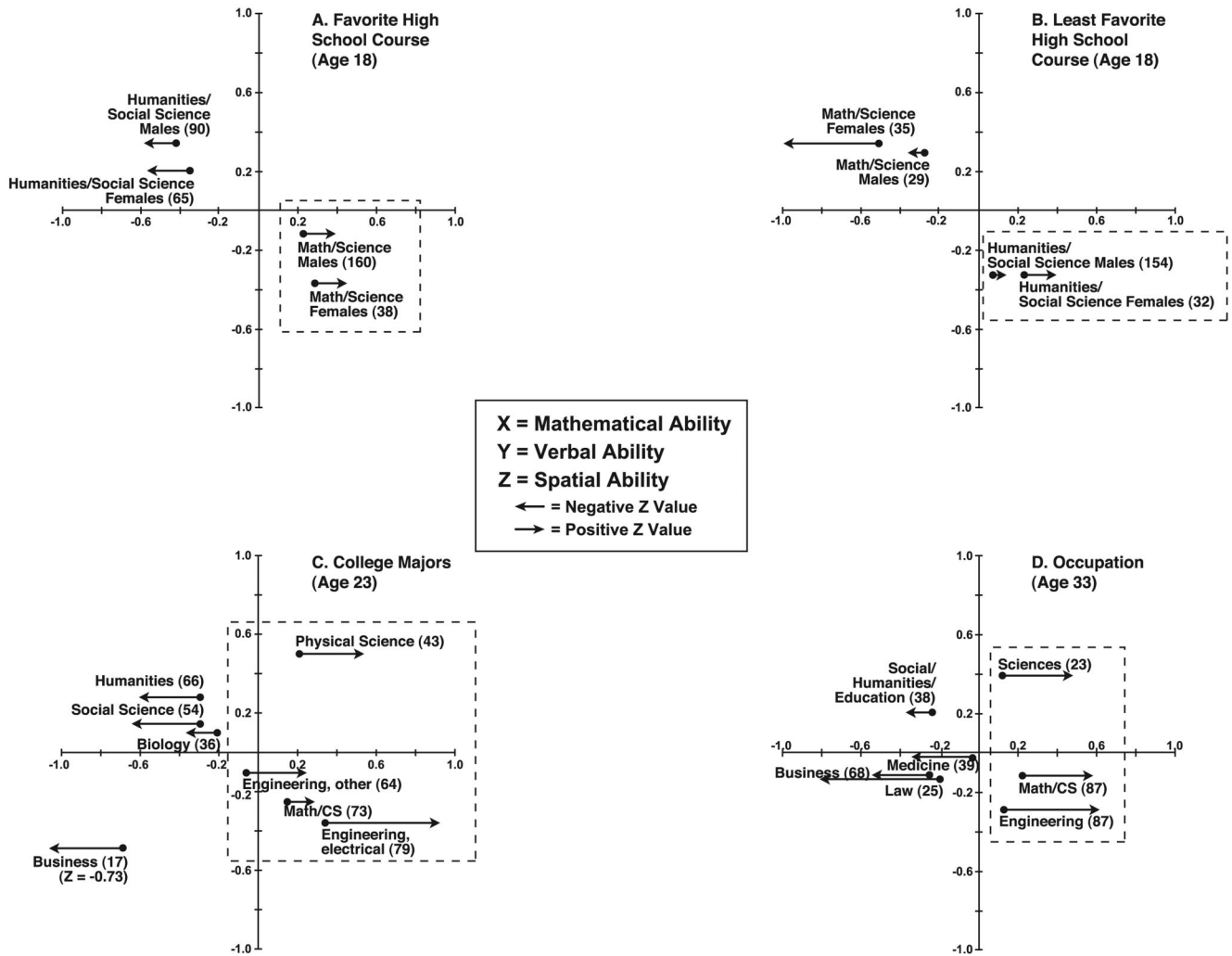
Figure 2. Shown are trivariate (X/Y/Z = Mathematical/Verbal/Spatial) means for (Panel A) favorite and (B) least favorite high school course at age 18, (C) college majors at age 23, and (D) occupation at age 33. Mathematical, verbal, and spatial ability are on the *x*-, *y*-, and *z*-axes respectively (arrows to the right indicate a positive *z* value; arrows to the left indicate a negative *z* value). Panels A and B are standardized within sex; Panels C and D are standardized across sexes. For Business in Panel C, note that the length of the arrow is actually *z* = −0.73. Figure adapted from Shea, Lubinski, & Benbow (2001). CS = computer science.

occupational criteria in STEM areas (where spatial ability is quite important) are highlighted by the rightward-pointing arrows for each of the STEM groups, which are contained within the dotted-line boxes in all four panels of Figure 2. The STEM groups were higher on spatial ability relative to the other groups; we refer to these groups later in the article. A key question is, has spatial ability been operating in this way in normative samples as well and, if so, for how long?

Replication Sample: Project TALENT

In this study, we formed a number of educational and occupational groupings using Project TALENT's 11-year follow-up data to reveal the extent to which spatial ability assessed in adolescence is a salient characteristic among individuals who subsequently go on to achieve educational and occupational

credentials in STEM. In addition, to ascertain the extent to which findings uncovered with Project TALENT mirror those in the Shea et al. (2001) study, we scrutinized the pattern similarity of specific abilities in both data sets. The question here is, when calibrated against STEM criteria, will both data sets reveal a consistent pattern (i.e., high mathematical and spatial ability and relatively lower verbal ability over multiple longitudinal time frames)? Moreover, the sample sizes available in Project TALENT allow us to examine whether higher levels of spatial ability differentiate people operating at more advanced educational levels within STEM. Finally, we identify the proportion of participants in the top 1% of spatial ability who are not in the top 1% on either mathematical or verbal ability and, hence, are lost by identification procedures restricted to mathematical or verbal ability; we examine the

educational and occupational outcomes of these students to understand better what kinds of students modern talent search procedures are failing to identify for advanced educational opportunities. For further, more detailed reading on the methodological approach of creating criterion groups based on educational and occupational credentials and, subsequently, examining the salient characteristics among members of each group at earlier time points for clues about the psychological antecedents giving rise to them, see Dawis (1992); Dawis and Lofquist (1984); and especially Humphreys et al. (1993) and references therein.

Method

Participants and Measures

Participants were drawn from the Project TALENT data bank, an ideal sample for our purposes here due to its comprehensiveness, size, and longitudinal time frame. Project TALENT's initial data collection in 1960 consisted of a stratified random sample of the nation's high school population (Flanagan et al., 1962). Students in the 9th through 12th grades were assessed on a wide range of tests and questionnaires over a 1-week period; the entire sample included roughly 50,000 males and 50,000 females per grade level, for a total N of approximately 400,000. Included in the tests were a number of measures designed to assess cognitive abilities (e.g., mathematical, verbal, and spatial ability), as well as information tests (on content areas including art, biology, engineering, journalism, and physics) and measures of attitudes, interests, and personality traits. Participants also completed a 398-item questionnaire on their lives (e.g., topics such as family, school, work, hobbies, and health). Tests and questionnaires were administered over a period of 1 week. These materials can be obtained through the American Institutes for Research, Palo Alto, California (see Flanagan et al., 1962, and Wise et al., 1979, for a thorough description of the range of tests and questionnaires administered).

Longitudinal Data

Project TALENT includes longitudinal data taken 1, 5, and 11 years after graduation from high school (Wise et al., 1979). For this study, we examined the 11-year follow-up data and focused on those who reported their highest degree received (a bachelor's, master's, or doctoral degree) and occupation.

Research Design

The conceptual framework used to form our ability measures stems from the hierarchical organization of cognitive abilities (Carroll, 1993). A cogent simplification of Carroll's model is the radex organization or scaling of cognitive abilities (Snow, Corno, & Jackson, 1996; Snow & Lohman, 1989). The radex organizes cognitive abilities around three content domains: quantitative/ numerical, spatial/pictorial, and verbal/linguistic (or mathematical, spatial, and verbal domains, respectively); the communality cutting across these three content domains distills the higher order construct of general intelligence (g). The latter denotes the sophistication of the intellectual repertoire. Figure 3 constitutes a visual representation of the radex, which is made up of an
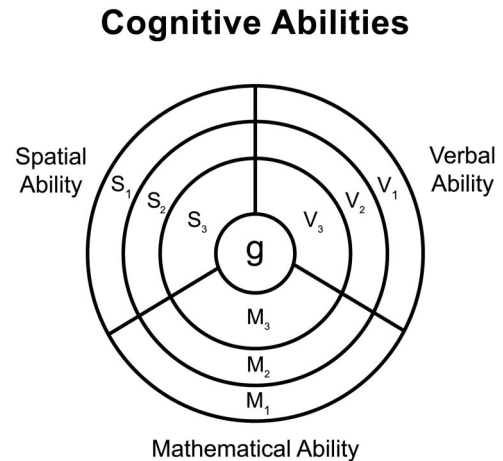
Cognitive Abilities

Figure 3. The radex of cognitive abilities organizes cognitive abilities around mathematical, spatial, and verbal content domains, with the higher order construct of *g*, or general intelligence, at the center representing the communality shared by these specific abilities.

infinite number of simplexes and circumplexes. An example of a simplex would be a continuum running from the centroid, or g, through S_3, S_2, and S_1, respectively. Along this simplex, the test content is spatial and the test complexity diminishes as one moves from the center to the periphery (e.g., a test located in S_3 would be more complex than a test located in S_1). Thus, for simplexes, test content remains comparable but complexity changes. An example of a circumplex is a circular band running through S_2, V_2, and M_2, respectively. Along this circumplex, the test content would vary, being spatial (reasoning with figures and shapes), verbal (reasoning with words), or mathematical (reasoning with numbers); however, the test complexity would remain comparable. Within the radex, tests varying in content and complexity can be found, and these two dimensions are necessary for locating a specific test in this space. The radex is a very efficient way of arranging the many different kinds of psychometric indicators of cognitive abilities. To the extent that measures covary with one another, they are close to one another in this two-dimensional space. Correspondingly, to the extent that measures do not covary with one another, they are distant from one another in this space (cf. Lubinski & Dawis, 1992, p. 8, for an empirical example). Thus, spatial ability, the focal construct under investigation, is distinguished from the more familiar constructs of mathematical and verbal ability in the context of a hierarchical illustration of the radex organization of cognitive abilities (see Figure 3).

Ability composites. We formed three ability composites with which to measure the three components found in the radex (Snow & Lohman, 1989; Wise et al., 1979): mathematical, spatial, and verbal ability. The Mathematical Composite consisted of four tests:

1. Mathematics Information (23 items measuring knowledge of math definitions and notation). A sample item might be "Which of these is an irrational number?"

2. Arithmetic Reasoning (16 items measuring the reasoning ability needed to solve basic arithmetic items). A sample

item might be "A man pays 4% sales tax on a chair. The tax is $6.00. How much did the chair cost?"

3. Introductory Mathematics (24 items measuring all forms of math knowledge taught through the 9th grade). A sample item might be "Suppose the sum of 2 two-digit numbers is a three-digit number. What is the first digit of the sum?"

4. Advanced Mathematics (14 items covering algebra, plane and solid geometry, probability, logic, logarithms, and

basic calculus). A sample item might be "Which of these equations has no real roots?"

To maximize construct validity (see below), we assigned the following weights based on scale variances and covariances to these constituents: Mathematical Composite $= 0.55 \times$ [Mathematical Information] $+ 1.0 \times$ [Arithmetic Reasoning] $+ 0.55 \times$ [Introductory Mathematics] $+ 1.0 \times$ [Advanced Mathematics].
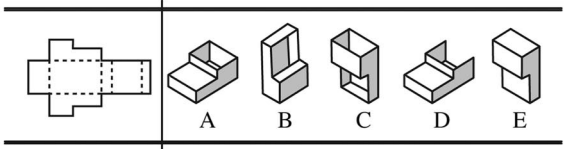
The Verbal Composite was composed of three measures:

1. Vocabulary (30 items that measure general knowledge of words). A sample item might be "Placate means" with answer choices following.

2. English Composite (113 items measuring capitalization, punctuation, spelling, usage, and effective expression). A sample item for covering usage might be "He _____ ready yet; A. isn't, B. ain't, or C. aren't."

3. Reading Comprehension (48 items measuring the comprehension of written text covering a broad range of topics). A sample item in this section would be similar to a typical reading comprehension item found on an exam such as the SAT.

Verbal Composite $= 2.5 \times$ [Vocabulary] $+ 1.0 \times$ [English Composite] $+ 1.25 \times$ [Reading Comprehension].
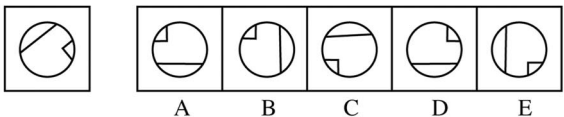
Finally, the Spatial Composite was composed of four measures (and because the focus of this study is on spatial ability, item types for each are illustrated in Figure 4):

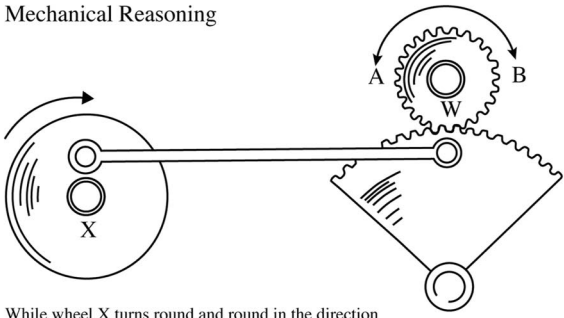1. Three-Dimensional Spatial Visualization (16 items measuring the ability to visualize two-dimensional fig-
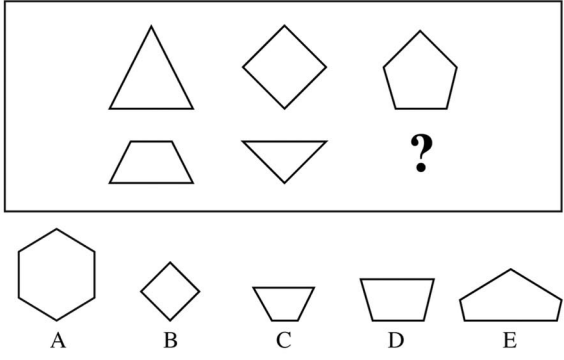
Figure 4. Three-dimensional spatial visualization. Each problem in this test has a drawing of a flat piece of metal at the left. At the right are shown five objects, only one of which might be made by folding the flat piece of metal along the dotted line. You are to pick out the one of these five objects which shows just how the piece of flat metal will look when it is folded at the dotted lines. When it is folded, no piece of metal overlaps any other piece or is enclosed inside the object.

Two-dimensional spatial visualization. In this test each problem has one drawing at the left and five similar drawings to the right of it, but only one of the five drawings on the right exactly matches the drawing at the left if you turn it around. The rest of the drawings are backwards even when they are turned around. For each problem in this test, choose the one drawing which, when turned around or rotated, is exactly like the basic drawing at the left.

Mechanical reasoning. This is a test of your ability to understand mechanical ideas. You will have some diagrams or pictures with questions about them. For each problem, read the question, study the picture above it, and mark the letter of the answer on your answer sheet.

Abstract reasoning. Each item in this test consists of a set of figures arranged in a pattern, formed according to certain rules. In each problem you are to decide what figure belongs where the question mark is in the pattern. To do this you have to figure out what the rule is according to which the drawings change, going from row to row, and what the rule is for the changes going from column to column. The items have different kinds of patterns and different rules by which the drawings change. The question mark in the lower right corner of each box shows where a figure is missing in the pattern. You are to decide which of the five figures (A, B, C, D, or E) under the pattern belongs where the question mark is.

ures after they had been folded into three-dimensional figures).

2. Two-Dimensional Spatial Visualization (24 items measuring the ability to visualize two-dimensional figures when they were rotated or flipped in a plane).

3. Mechanical Reasoning (20 items measuring the ability to deduce relationships between gears, pulleys, and springs as well as knowledge of the effects of basic physical forces, such as gravity).

4. Abstract Reasoning (15 items constituting a nonverbal measure of finding logical relationships in sophisticated figure patterns).

Spatial Composite = 3.0 × [3-D Spatial Visualization] + 1.0 × [2-D Spatial Visualization] + 1.5 × [Mechanical Reasoning] + 2.0 × [Abstract Reasoning].

The above weights were derived by Humphreys to form composites that mirror the SAT-M and SAT-V and the location of spatial ability within the context of the hierarchical organization of cognitive abilities (see the radex in Figure 3), and these composites have been used extensively in other research (Gohm et al., 1998; Humphreys et al., 1993; Lubinski & Humphreys, 1990a, 1990b, 1996). The intercorrelations of these composites for the 9th-grade cohort were .61, .59, and .76 for mathematical–spatial, verbal–spatial, and mathematical–verbal, respectively. In the current study, we added the English Composite and Advanced Mathematics to their respective composites initially derived by Humphreys to augment the ceiling of each scale, but this modification changed the intercorrelations of these three composites by an average of only .01 correlational units for each sex across all four cohorts. Hence, their conceptual equivalency and empirical interchangeability were preserved. Humphreys has estimated that the reliabilities of these, or very similar, composites are approximately .90 (Humphreys, 1991). These estimates were based on conservative estimates of parallel form reliabilities of the components.

Design. Participants were included if they had complete ability data at Time 1: (*n*s: 9th grade, males = 47,440, females = 47,496; 10th grade, males = 46,112, females = 45,199; 11th grade, males = 41,766, females = 43,751; 12th grade, males = 36,375, females = 38,526). Would the educational and occupational group membership of these participants, assessed 11 years after their high school graduation, retrospectively isolate distinctive ability profiles based on their adolescent assessments? If so, and if these findings mirrored those uncovered in Shea et al. (2001) over a 20-year interval and nonoverlapping time frame (see Figure 2), these corresponding function forms would constitute a *constructive replication* (Lykken, 1968, 1991).

In this context, it is worth mentioning Meehl's (1978) point that in the early stages of theory construction, function form is often more important than statistical significance (see also Steen, 1988). For example, the patterns of specific abilities in all four panels of Figure 2 reveal consistent function forms. Across all four panels, meaningful STEM outcomes are found in roughly the same location as a function of the three specific abilities; over all time points, they reveal the same pattern (see the groups

in the dotted-line boxes). That is, the same function form or pattern of specific abilities distinguishes the STEM groups from the other criterion groupings over these multiple time points (high school, college, and occupation). The precise location of the points on each panel is not as critical as the overall pattern formed by the specific abilities over time. They appear to be operating in the same way, and the pattern maintains its function form. We hypothesized that similar function forms would be found using Project TALENT (a different cohort and nonoverlapping time frame) and that these patterns would be of sufficient magnitude to be of substantive interest to psychological practitioners, applied researchers, and theoreticians interested in educational readiness and adult achievement. Finally, although both function form and statistical significance are evaluated here, the former is more central because, given the sample sizes, virtually all group contrasts will manifest statistically significant differences on the specific abilities under analysis.

We made the following hypotheses:

1. The pattern or function form uncovered from the participants in SMPY on the three specific abilities will be mirrored by those in Project TALENT when calibrated against conceptually meaningful educational and occupational criterion groupings.

2. The importance of spatial ability will increase as a function of more conceptually demanding STEM criteria (e.g., advanced educational degrees in STEM: bachelors, masters, and doctorates).

3. To the extent possible, findings taken from the Graduate Record Exam will mirror those of Project TALENT and SMPY.

4. An appreciable percentage of young adolescents with talent for STEM and other domains in which spatial ability is placed at a premium are missed by contemporary talent searches and current selection procedures for STEM careers.

To begin to examine this series of hypotheses, after selecting participants in each cohort as a function of complete ability and group membership data, we computed *z* scores for all three specific abilities within cohort and, then, over all four cohorts. Data for each criterion group (within highest degree and occupation) were aggregated. We then plotted each group's mean *z* score for all three specific abilities (see Figure 5). Appendix A includes the respective sample sizes broken down by sex for each degree and occupation included in the groups plotted over all four panels of Figure 5.[1]

---

[1] We conducted analyses for males and females separately within each grade level (9th, 10th, 11th, and 12th) and found that the pattern was strikingly similar.
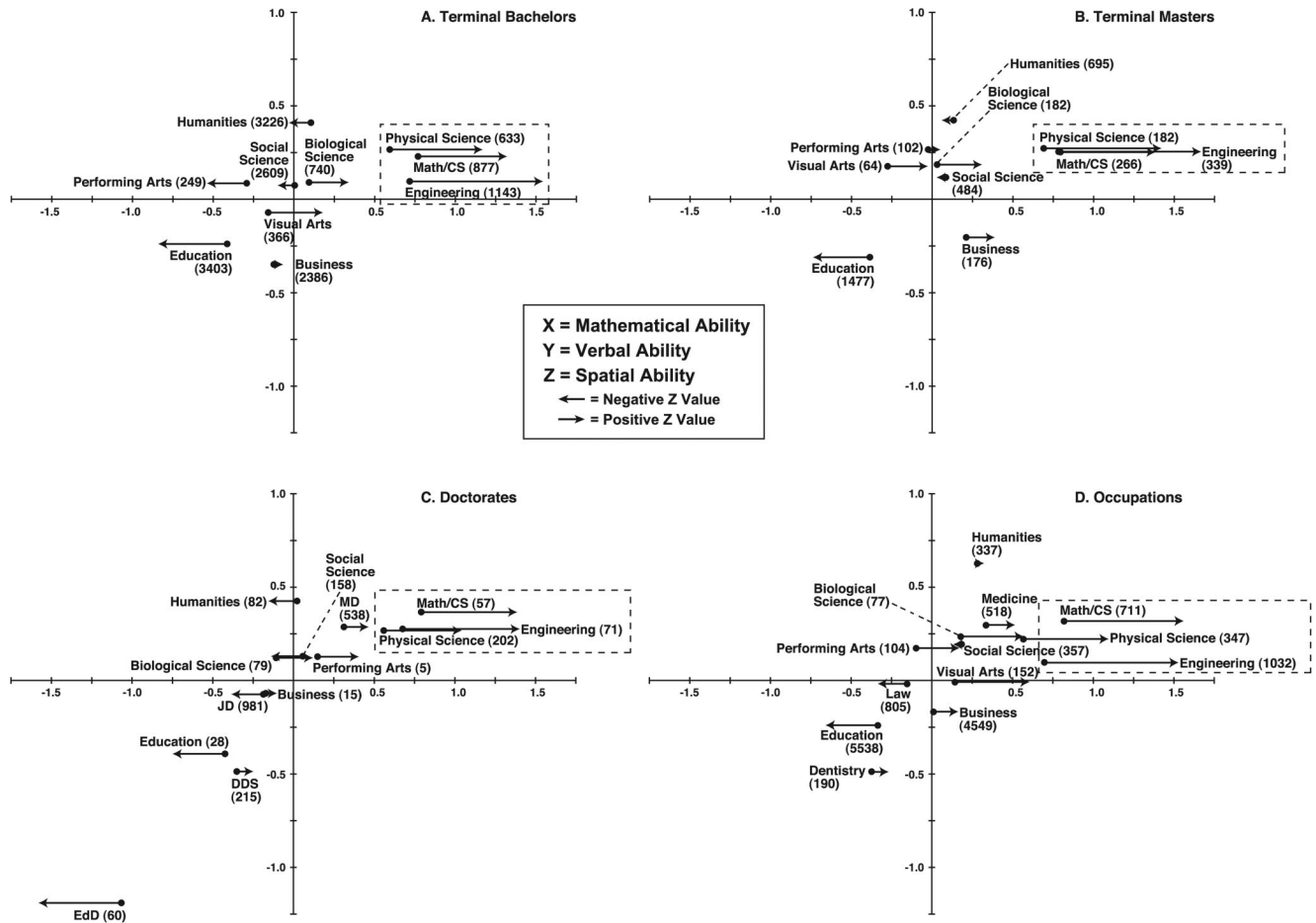
Figure 5. Trivariate means for (Panel A) bachelors, (B) masters, (C) doctorates, and (D) occupations of those individuals whose data were included in Panels A, B, and C. Panels A through D are standardized across sexes. Mathematical ability is on the *x*-axis, and verbal ability is on the *y*-axis; an arrow from each group mean indicates either positive (to the right) or negative (to the left) spatial ability. Breakdowns by sex are reported in Appendix B. The standard error of the mean for $n = 500$ was 0.04 and for $n = 1,000$ was 0.03. Data are from Project TALENT. CS= computer science.

Results

Ability Pattern

A number of the groups included in Figure 5 are conceptually similar to those in the Shea et al. (2001) panels found in Figure 2. Those that are not were included to provide a more detailed context for how these ability composites operate within a more broadly defined educational–occupational criterion space. Figure 5 includes those individuals who reported at the 11-year follow-up that their highest degree received was a BA or BS (Panel A; males = 8,446, females = 7,186), a MA or MS (Panel B; males = 2,383, females = 1,584), or a doctorate (Panel C; males = 2,293, females = 198). In addition to PhDs, which correspond to the degrees found in Panels A and B, Panel C contains MDs, JDs, DDSs, and EdDs for comprehensiveness. Panels A, B, and C included nonoverlapping groups of participants, as only highest degrees received were plotted. Panel D (males = 10,389, females = 4,328) included all participants

who reported a degree as shown in Panels A, B, and C and who also reported an occupation 11 years after high school graduation. Each graph parallels the Figure 2 template; mathematical (*x*-axis), verbal (*y*-axis), and spatial (*z*-axis) abilities are plotted in standard deviation units. Sample sizes are next to each group in parentheses.

One can see in each panel a general ability, or *g,* gradient (driven by the communality running through spatial, mathematical, and verbal ability), extending from the lower left quadrant to the upper right quadrant. It is also evident that the patterns of specific abilities (within each group and across groups) shown in the four panels are strikingly similar. As an example, within the dotted-line boxes in each panel of Figure 5 are the STEM groups (as in Figure 2). Examination of the physical science group in Panel C shows that this group has a positive *z*-score value (relative to the mean) on mathematical, verbal, and spatial ability. It is important here to note that in both Figures 2 and 5 the STEM groups all have rightward-pointing arrows, which indicate higher spatial ability

(relative to the other groups in each panel). The rightward-pointing arrow in Panel C for the physical sciences is 0.45 standard deviation units greater than the mean, whereas the humanities grouping is −0.15 standard deviation units below the mean; therefore, these two groups are 0.60 standard deviations apart on spatial ability. This constitutes one of many examples of an important constructive replication of function form or pattern across Figures 2 and 5 for spatial ability, which is central to our thesis.[2] It is important to note here that the locations of mathematical (*x*-axis) and verbal (*y*-axis) abilities for some groups are offset somewhat over Panels A through D in Figure 5 in comparison to those of Figure 2, but for good reason.

Dawes (1975) and Sackett and Yang (2000) have discussed how structural relationships among measures can change when samples under analysis are selected with predictors that go into subsequent analyses.[3] Participants summarized in Figure 2 were selected using the SAT, whereas Figure 5 participants were a stratified random sample of the nation's high school population who subsequently went on to earn advanced educational credentials. As noted before, neither cohort was selected on spatial ability. It is striking, therefore, how clear-cut the findings are for corresponding groups on spatial ability, especially in STEM domains. Moreover, other domains, such as biology and the visual arts, also appear to draw on spatial ability. For other, more general consistencies found in these data, Appendix B provides the overall level of general and specific abilities that these groups manifested at Time 1.[4]

In Figure 6, we extend these analyses in part to another contemporary sample by plotting bivariate (X/Y = Mathematical/Verbal) means for the mathematical and verbal ability composites from Project TALENT for participants who later went on to secure graduate degrees (black circles). The bivariate means for these participants are connected with lines to the bivariate means of contemporary graduate students on the basis of corresponding mathematical and verbal measures (white circles) on the Graduate Record Examination (GRE). Figure 6 is similar to Figures 2 and 5 in that it includes mathematical ability on the *x*-axis and verbal ability on the *y*-axis, but it is different in that it does not include spatial ability (rightward- and leftward-pointing arrows), inasmuch as the GRE does not include a spatial measure. Again, mathematical ability is a salient attribute of students seeking to develop STEM expertise. As described in the caption, each GRE grouping represents thousands of prospective graduate students. Although the GRE does not assess spatial ability, given the consistencies between the GRE and Project TALENT's mathematical and verbal ability scales and the well-known longitudinal consistencies of the covariance structures between measures of these constructs (Carroll, 1993; Johnson & Bouchard, 2007a, 2007b; Lubinski, 2004; Snow et al., 1996), it would be surprising if modern spatial ability assessments did not uncover patterns consistent with the other longitudinal findings (see, e.g., Figures 2 and 5). Essentially, we anticipated that the spatial ability arrows on the *z*-axis for the GRE data, if plotted, would reflect those found in Figures 2 and 5. In Figure 6, graduate degrees in the humanities are high on the *y*-axis, and a salient cluster of graduate degrees in STEM are located far to the right on the *x*-axis (engineering, math/computer science, and physical science); business and education also demonstrate a co-occurrence of location across both data sets in the space defined by these dimensions. Consistent locations are therefore found over a 40-year interval.

Spatial Ability Level

With respect to overall level of ability, the likelihood of earning an advanced degree in STEM as a function of spatial ability is depicted in Figure 7. Using 11-year follow-up data from Project TALENT, we classified the subset of participants with STEM degrees into three groups (as a function of their highest terminal degree): bachelor's, master's, or PhD. This was done within each cohort separately, and then findings from all four cohorts were aggregated. Finally, we plotted the proportion of each degree within each stanine based on spatial ability stanine in high school. It becomes clear from these findings that spatial ability plays an important role in achieving advanced educational credentials in STEM. From an epidemiological point of view (Lubinski & Humphreys, 1996, 1997), the likelihood or promise of earning an advanced degree in STEM areas increases as a function of spatial ability. These findings are clear: 45% of all those holding STEM PhDs were in Stanine 9 (or within the top 4%) on spatial ability 11+ years earlier, and nearly 90% were in Stanine 7 or above. That is, less than 10% of those holding STEM PhDs were below the top quartile in spatial ability during adolescence. In comparison to the 45% of STEM PhDs in Stanine 9, for example, about 30% of those holding STEM terminal master's degrees and 25% of those holding STEM terminal bachelor's degrees were in Stanine 9, or the top 4% of spatial ability. We can conclude that the importance of spatial ability for STEM increases as a function of successively more advanced educational credentials.

Finally, is there a way to determine the extent to which modern talent searches miss high-potential students gifted in spatial ability? Some summer residential programs for talented youths require scores in the top 1% on either verbal or mathematical ability measures to ensure readiness to take advantage of the fast-paced learning demands

---

[2] For all four panels in Figure 5, we also conducted the following series of incremental validity analyses. For each of the three terminal degrees and occupations, we dummy coded each STEM cluster as 1 and the remainder of the groups as 0 (and we utilized this as a criterion variable). We then ran multiple regression analyses for each panel by first entering mathematical ability and verbal ability and then determining the incremental validity, or multiple-R^2 increment, for spatial ability in predicting this dichotomous variable (STEM, non-STEM). The incremental validity of spatial ability over mathematical ability and verbal ability for all four panels was statistically significant, as anticipated, and the multiple-R^2 increment averaged .04 (accounting for an additional 4% of criterion variance).

[3] The relationship found in nature between two variables actually can be inverted (a positive covariance can become negative) when selection occurs on a third variable. For example, among undergraduates applying to graduate school, their composite Graduate Record Exam (GRE) score is positively correlated with their undergraduate grade point average (GPA) but is negatively correlated among those selected within a particular school. The reason is that a low GRE composite can be compensated for by a high GPA, and the inverse is true for low GPAs. But for graduate school, the low-GRE, low-GPA students tend not to be selected; this removal of the southwest quadrant of the fourfold table (GRE/GPA, High/Low) switches a positive covariance to a negative one.

[4] In a related vein, as the 9th grade of Project TALENT is closest to the SMPY talent search population in time of initial testing, we conducted an analysis to determine how similar all four cohorts were to the 9th-grade cohort alone. The average difference between all four grades and the 9th-grade sample for the respective correlations for mathematical, spatial, and verbal abilities was less than 0.03 correlational units.
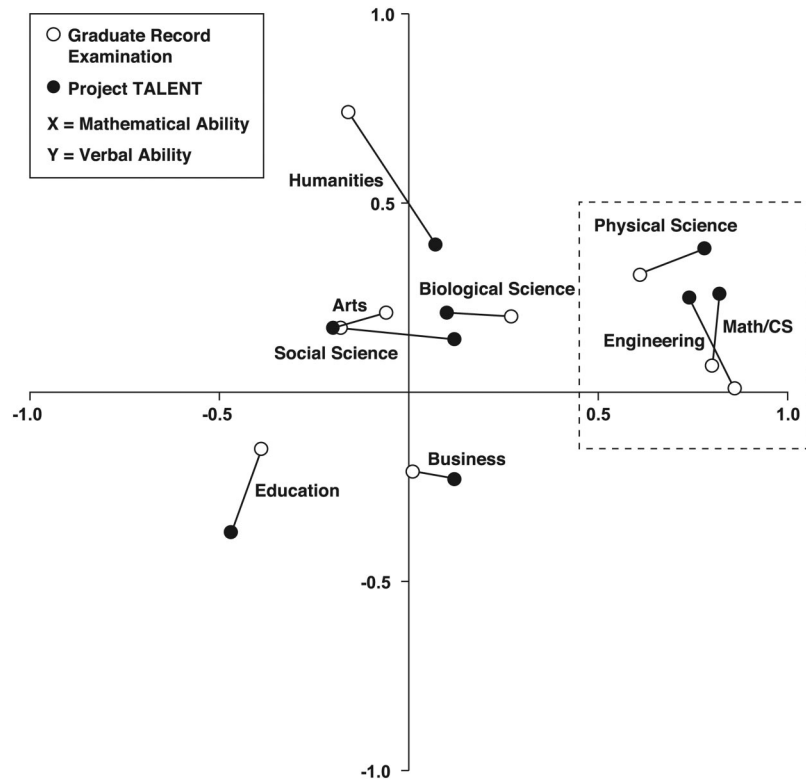
Figure 6. Data on the Graduate Record Examination (GRE) taken on individuals in the respective educational groups tested between July 1, 2002, and June 30, 2005, are graphed alongside commensurate data from Project TALENT. For each group, *z* scores on mathematical (*x*-axis) and verbal (*y*-axis) abilities respectively are plotted (standardized within the groups represented). White circles = GRE data. Black circles = Project TALENT data. A line was drawn connecting the two data points for each group to illustrate the distances between points of the same field. The *z* scores for each group were computed by taking the difference between the group mean and the overall mean for each subtest and dividing by the population standard deviation of that subtest. The total number of those taking the GRE for each subtest for these data was 1,245,872 for GRE-Mathematical (GRE-M) and 1,245,878 for GRE-Verbal (GRE-V). The respective groups were chosen to mirror the ones in Figure 2 and were as follows (with *n*s for GRE-V and GRE-M, respectively): engineering (56,368, 56,374); physical science (22,487, 22,485): chemistry, earth, atmospheric, and marine sciences, and physics and astronomy; math/computer science (33,107, 33,108): computer and information sciences, mathematical sciences; biological science (37,579, 37,576); humanities (37,468, 37,435): English language and literature, foreign languages and literatures, history, philosophy, and religion and theory; social science (101,085, 101,064); arts (20,040, 20,057): architecture and environmental design, art history, theory, and criticism and arts, performance and studio; business (8,357, 8,357); education (43,844, 43,835). Project TALENT data (PT-M, PT-V) were analyzed within MAs, MSs, and PhDs specifically to best mirror the GRE data. Correlations between the means for the respective educational groups were computed between GRE-M and PT-M ($r = .93$, $p < .01$), GRE-V and PT-V ($r = .77$, $p < .05$), and GRE-M + V and PT-M + V ($r = .96$, $p < .01$). The average difference across all three methods of comparison (i.e., correlations GRE-M minus PT-M, GRE-V minus PT-V, and GRE-M + V minus PT-M + V) and major groupings was less than the absolute value of 0.04, 0.02, and 0.02, respectively. The standard error of the mean for $n = 500$ was 0.04 and for $n = 1,000$ was 0.03. GRE data were taken from http://www.ets.org/Media/Tests/GRE/pdf/5_01738_table_4.pdf and http://www.ets.org/Media/Tests/GRE/pdf/4_01738_table_1a.pdf

in their programs, and some even require scores in the top 0.5% (Benbow & Stanley, 1996; Colangelo et al., 2004). Thus, there exists another question that Project TALENT can answer: How many spatially gifted students are missed for such programs by current talent search practices, which focus only on mathematically and verbally talented youths? Within the three ability composites assembled for this study, 70% of the top 1% in spatial ability did not make the cut for the top 1% on either the math or the verbal composite; yet, these

individuals are highly talented in spatial ability. Figure 8 presents data on the educational and occupational outcomes of this 70% in terms of their credentials in STEM domains (top panel) and the visual arts (bottom panel). The latter group was added to highlight the longstanding recognition of the importance of spatial ability for many of the creative arts. The black bars show the base rates for these outcomes in Project TALENT; the overall bars (black + gray) represent those in the top 1% on the Spatial Composite who were not in the top
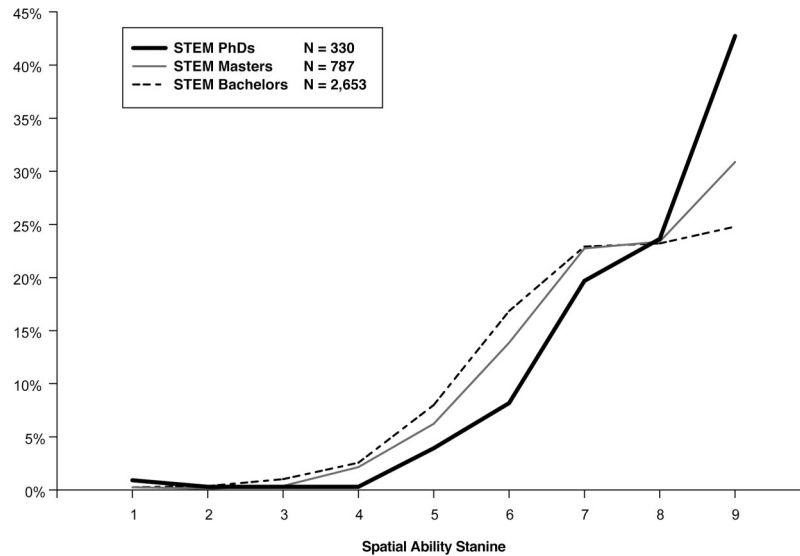
Figure 7. This figure includes the proportion of each degree group (bachelors, masters, and PHDs) as a function of spatial ability. Along the *x*-axis are the spatial ability stanines (numbered 1 through 9). STEM = science, technology, engineering, and mathematics.

1% on either the Math or Verbal Composites. This potential, currently being missed, constitutes a rather sizable pool of untapped talent. Among those in the top 1% in spatial ability but not in the top 1% in mathematical or verbal ability, a large proportion earned STEM and visual arts degrees and entered STEM and visual arts occupations well beyond base rate expectations.

Discussion

Longitudinal findings uncovered in this study combined with results of earlier investigations (Super & Bachrach, 1957) and recent longitudinal findings on intellectually precocious youths (Shea et al., 2001; Webb et al., 2007) suggest at least three generalizations: First, spatial ability is a salient psychological characteristic among adolescents who subsequently go on to achieve advanced educational and occupational credentials in STEM. Second, spatial ability plays a critical role in structuring educational and occupational outcomes in the general population as well as among intellectually talented individuals. Third, contemporary talent searches miss many intellectually talented students by restricting selection criteria to mathematical and verbal ability measures.[5]

Given the body of evidence now available and the fresh empirical findings presented here on thousands of high school students tracked 11 years following their high school graduation, sufficient support has accrued to demonstrate that the importance of spatial ability in STEM domains has been operating for several decades. Just as F. L. Schmidt and Hunter (1998) concluded in their 85-year review of the role that general intelligence plays in the world of work ("more research is not needed"), we conclude that enough empirical evidence has accrued to register another rare example of a solid empirical generalization within the human psychological sciences. This does not mean, however, that other research is not needed. The kind of research that is needed now is in how to utilize

spatial ability for student selection, instruction, and curriculum design and in how to refine educational interventions and procedures on the basis of individual differences in spatial ability (Corno et al., 2002; Lubinski, 2004, pp. 105–106).

In addition, measures of spatial ability should be incorporated into models of educational and occupational development to ascertain the role spatial ability plays relative to other abilities and relevant nonintellectual determinants (Lubinski & Benbow, 2000, 2006). Given the evidence presented here, psychological modeling of STEM outcomes must incorporate spatial ability to avoid being incomplete or underdetermined (Lubinski, 2000; Lubinski & Humphreys, 1997). This is particularly true among those who go on to develop especially high levels of STEM expertise (cf. Figures 2, 5, 7, and Appendix B).

Furthermore, expanding admissions criteria for talent searches currently focused on identifying intellectually talented youths

---

[5] There have been some discussions in visible outlets and based on very small samples that socioeconomic status (SES) moderates the sex difference in spatial ability (Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). Levine et al. has been cited by a number of recent investigations (Alexander & Son, 2007; Bergemann et al., 2008; Chabris & Glickman, 2006; Ehrlich, Levine, & Goldin-Meadow, 2006; Hackman & Farah, 2009; Newcombe & Uttal, 2006; Noble, McCandliss, & Farah, 2007; Penner & Paret, 2008; Silverman, Choi, & Peters, 2007) as documenting this relationship. As Newcombe and Uttal (2006) further generalize, "We need to delineate why and how some of the core abilities that all humans have come to be developed to different degrees in ways that depend on interactions of SES and gender" (p. 395). We conducted an analysis with our spatial ability composite (see Appendix C), in which we divided the Project TALENT SES variable into four quartiles and examined by cohort and by sex at each level of SES the hypothesis that SES moderates the sex difference in spatial ability ($n \approx 10,000$ in each cell). As can be seen in Appendix C, it is simply not the case that SES moderates the sex difference in spatial ability.
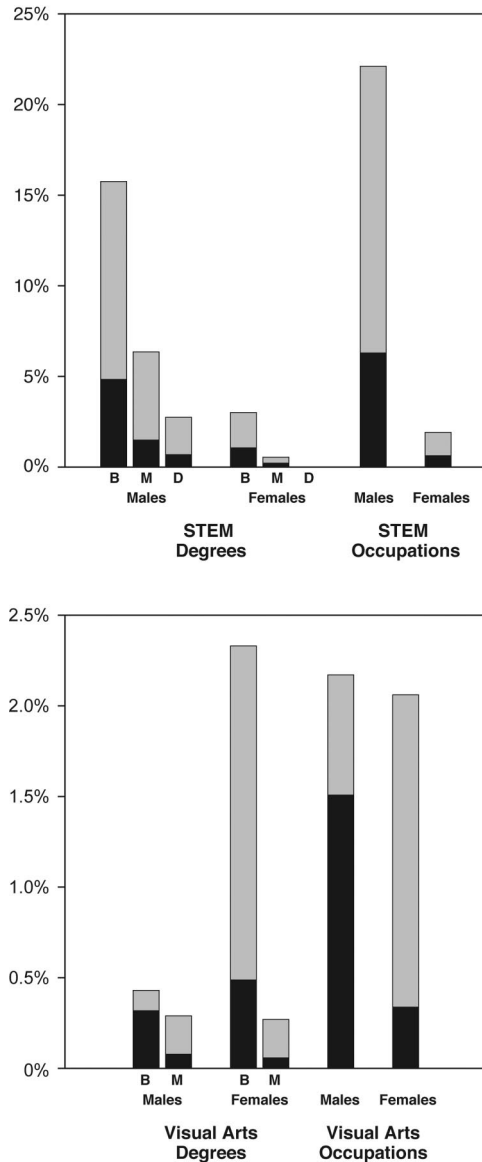
Figure 8. The top panel includes (left) the proportion of the top 1% in spatial ability who were not in the top 1% in mathematical or verbal ability who earned STEM degrees and (right) occupations broken down by males and females, respectively. The bottom panel includes the proportion of this population who earned visual arts degrees and worked in related occupations. The black bars indicate the base rate in Project TALENT for the respective grouping. B = bachelor's degrees; M = master's degrees; D = doctorate degrees; STEM = science, technology, engineering, and mathematics.

solely on the basis of scores on mathematical and verbal measures should be considered, as Snow (1999) suggested. For example, in Project TALENT, over half of participants in the top 1% on the Spatial Composite were below the top 3% cut on both the Mathematical and Verbal Composites, and, thus, they would not be invited to participate in modern talent searches. Moreover, there is reason to believe that the educational needs of spatially talented youths are more unmet than those of mathematically or verbally talented youths, because the typical middle and high school cur-

riculum has many more opportunities for developing mathematical and verbal ability than spatial ability (Colangelo et al., 2004; Lohman, 2005).[6]

Ackerman (1996; Ackerman & Heggestad, 1997) has shown that all three of the specific abilities examined here (mathematical, spatial, and verbal) have distinctive external correlational profiles with respect to conventional measures of interests, personality, and values. For example, verbal ability tends to covary positively with interests in the humanities and helping people and to covary negatively with interests in engineering and technical pursuits. The opposite is true for spatial ability. An examination of the intercorrelations of Project TALENT's ability and interest measures reveals that these trait clusters, too, have been observed for decades in normative samples (Shaycoft, 1967) and, thus, must be seen as relatively stable. Moreover, these patterns of covariation have been replicated with intellectually talented youths (D. B. Schmidt, Lubinski, & Benbow, 1998) and have emerged as salient weights in discriminant function analyses in the prediction of STEM criteria (Achter, Lubinski, Benbow, & Eftekhari-Sanjani, 1999; Wai, Lubinski, & Benbow, 2005; Webb et al., 2007). Therefore, the motivational proclivities for students selected on the basis of mathematical versus spatial versus verbal ability should be expected to differ. That is, intellectually talented students selected by extreme cutting scores on measures of mathematical versus spatial versus verbal ability should be expected to have different interest patterns as well as differential preferences for linguistic, quantitative, and nonverbal ideation or contrasting modes of learning and thought (Corno et al., 2002).

Just as mathematically and verbally talented students have profited for decades by talent searches that identify students especially able at verbal and mathematical reasoning and the provision of tailored, developmentally appropriate curriculum aligned to their precocious rates of learning (or reasoning with linguistic and numerical symbols, respectively), students talented in spatial ability are likely to profit from identification procedures utilizing measures of spatial ability followed by opportunities for developmentally appropriate curriculum involving their preferred mode of thought (reasoning with forms or shapes). Experimentation with accelerative and rigorous learning opportunities in architecture, engineering, robotics, and the physical sciences appear to be particularly warranted in order to nurture their form of talent.[7]

Finally, sex differences in relative levels of interests are important to take into consideration. Although the covariance structure of specific abilities and interests is comparable for males and for females, the sexes display mean differences in a number of interests; for instance, spatially talented females tend to be more interested in artistic pursuits than are spatially talented males and

---

[6] For further and more detailed reading on measures of spatial ability and their conceptual underpinnings, see Corno et al. (2002); Eliot (1987); Eliot and Smith (1983); Lohman (1988, 1994a, 1994b, 1996, 2005); and Vandenberg and Kuse (1978). For more historical accounts, which have acknowledged the importance of spatial ability for technical trades and professions, see Paterson, Elliott, Anderson, and Toops (1930); Smith (1964); and Vernon (1961).

[7] This move would also foster conditions for adding value to longitudinal models of creativity currently restricted to mathematical and verbal reasoning abilities (Park, Lubinski, & Benbow, 2007, 2008).

the inverse is true for engineering and mechanical activities (Lubinski & Benbow, 2006; D. B. Schmidt et al., 1998). These mean sex differences in interests correspond to findings, shown in Figure 8, that spatially talented females were more likely than similarly talented males to pursue artistic domains. These proclivities can and do change over time, but relative levels of interests (and competing interests) are always important to take into account (Geary, 1998, 2005; Gottfredson, 2003, 2005).

Cumulative Psychological Knowledge

Collectively, the findings reported here, when combined with Super and Bachrach's (1957) NSF report and linked to modern research on talent search participants (Shea et al., 2001; Webb et al., 2007), tell a cohesive story about the longitudinal stability of spatial ability and its psychological import (see Figures 2, 5, 7, and Appendix B). For decades, spatial ability has emerged as a salient psychological characteristic among young adolescents who go on to develop expertise in STEM domains (see Figure 7).

This fact is important for more general considerations, because in psychology the lack of cumulative knowledge upon which to build theory and practice is often bemoaned. Cronbach (1975) has discussed the short "half-life" of empirical generalizations in the social sciences (i.e., how quickly they decay) and wrote, "The trouble, as I see it, is that we cannot store up generalizations and constructs for ultimate assembly into a network" (p. 123). Similarly, Meehl (1978) has observed that the "soft areas of psychology lack the cumulative character of scientific knowledge" (p. 806). Leaders in industrial (Dunnette, 1966) and clinical psychology (Dawes, 1994) have echoed these remarks. The current study offers an example of how the human psychological sciences can generate cumulative knowledge. Teaming constructive replication with longitudinal inquiry appears to be a compelling way to achieve cumulative psychological knowledge by revealing consistent function forms both across and within cohorts over protracted intervals.

Conclusion

As I. M. Smith (1964) stated so well 45 years ago,

> The qualities which make for greatness in scientists and engineers are of a different kind; ability to think abstractly and analytically together with skill in visualizing spatial relations in two or three dimensions, . . . All these qualities, which are vitally important in almost all branches of science and engineering, are measured by appropriate tests of spatial ability. (p. 300)

Spatial ability's robust influence on STEM domains has been supported in this article through the presentation of findings that link decades of longitudinal research. Collectively, the studies presented here constitute a series of constructive replications revealing similarities in function form and pattern across time (Meehl, 1978, 1990; Steen, 1988); therefore, an empirical generalization may be ventured on the importance of spatial ability in scientific and technical domains. In addition, individuals who are high in spatial ability but not as exceptional in mathematical or verbal abilities constitute an untapped pool of talent for STEM domains. Currently, more research is needed on how to effectively structure educational opportunities to serve students talented in

spatial ability. Such efforts, if successful, will contribute to the urgent social need of effectively identifying and developing scientific and technical talent for the information age.

References

Achter, J. A., Lubinski, D., Benbow, C. P., & Eftekhari-Sanjani, H. (1999). Assessing vocational preferences among gifted adolescents adds incremental validity to abilities: A discriminant analysis of educational outcomes over a 10-year interval. *Journal of Educational Psychology, 91,* 777–786.

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22,* 227–257.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121,* 218–245.

Alexander, G. M., & Son, T. (2007). Androgens and eye movements in women and men during a test of mental rotation ability. *Hormones and Behavior, 52,* 197–204.

American Competitiveness Initiative. (2006). *American Competitiveness Initiative: Leading the world in innovation.* Washington, DC: Domestic Policy Council Office of Science and Technology.

Benbow, C. P., & Stanley, J. C. (1996). Inequity in equity: How "equity" can lead to inequity for high-potential students. *Psychology, Public Policy, and Law, 2,* 249–292.

Bergemann, N., Parzer, P., Kaiser, D., Maier-Braunleder, S., Mundt, C., & Klier, C. (2008). Testosterone and gonadotropins but not estrogen associated with spatial ability in women suffering from schizophrenia: A double-blind, placebo-controlled study. *Psychoneuroendocrinology, 33,* 507–516.

Bingham, W. V. (1937). *Aptitudes and aptitude testing.* New York: Harper.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Chabris, C. F., & Glickman, M. E. (2006). Sex differences in intellectual performance: Analysis of a large cohort of competitive chess players. *Psychological Science, 17,* 1009–1107.

Colangelo, N., Assouline, S. G., & Gross, M. U. M. (Eds.). (2004). *A nation deceived: How schools hold back America's brightest students.* Iowa City: University of Iowa.

Corno, L., Cronbach, L. J., Kupermintz, H., Lohman, D. F., Mandinach, D., Porteus, A. W., et al. (2002). *Remaking the concept of aptitude: Extending the legacy of Richard E. Snow.* Mahwah, NJ: Erlbaum.

Cronbach, L. J. (1957). Two disciplines of scientific psychology. *American Psychologist, 12,* 671–684.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30,* 116–127.

Dawes, R. M. (1975, February 28). Graduate admission variables and future success. *Science, 187,* 721–723.

Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on a myth.* New York: Free Press.

Dawis, R. V. (1992). The individual differences tradition in counseling psychology. *Journal of Counseling Psychology, 39,* 7–19.

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment.* Minneapolis: University of Minnesota Press.

Dunnette, M. D. (1966). Fads, fashions, and folderol in psychology. *American Psychologist, 21,* 343–352.

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42,* 1259–1268.

Eliot, J. C. (1987). Models of psychological space: Psychometric, developmental, and experimental approaches. New York: Springer-Verlag.

Eliot, J. C., & Smith, I. M. (1983). *An international dictionary of spatial tests.* Windsor, England: NFER-Nelson.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorman, W. A., Orr, D. B., & Goldberg, I. (1962). *Design for a study of American youth.* Boston: Houghton Mifflin.

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century.* New York: Farrar, Straus & Giroux.

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

Geary, D. C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence.* Washington, DC: American Psychological Association.

Gohm, C. L., Humphreys, L. G., & Yao, G. (1998). Underachievement among spatially gifted students. *American Educational Research Journal, 35,* 515–531.

Gottfredson, L. S. (2003). The challenge and promise of cognitive career assessment. *Journal of Career Assessment, 11,* 115–135.

Gottfredson, L. S. (2005). Using Gottfredson's theory of circumscription and compromise in career guidance and counseling. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: Putting theory and research to work* (pp. 71–100). New York: Wiley.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences, 13,* 65–73.

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121–169). New York: Cambridge University Press.

Humphreys, L. G. (1991). Some unconventional analyses of resemblance coefficients for male and female monozygotic and dizygotic twins. In D. Cicchette & W. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul Everett Meehl* (pp. 158–187). Minneapolis: University of Minnesota Press.

Humphreys, L. G., & Lubinski, D. (1996). Brief history and psychological significance of assessing spatial visualization. In C. P. Benbow & D. Lubinski (Eds.), *Intellectual talent: Psychometric and social issues* (pp. 116–140). Baltimore: Johns Hopkins University Press.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology, 78,* 250–261.

Johnson, W., & Bouchard, T. J. (2007a). Sex differences in mental abilities: *g* masks the dimensions on which they lie. *Intelligence, 35,* 23–39.

Johnson, W., & Bouchard, T. J. (2007b). Sex differences in mental ability: A proposed means to link them to brain structure and function. *Intelligence, 35,* 197–209.

Keating, D. P., & Stanley, J. S. (1972). Extreme measures for the mathematically gifted in mathematics and science. *Educational Researcher, 1,* 3–7.

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16,* 841–845.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 181–248). Hillsdale, NJ: Erlbaum.

Lohman, D. F. (1994a). Spatial ability. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (Vol. 2, pp. 1000–1007). New York: Macmillan.

Lohman, D. F. (1994b). Spatially gifted, verbally inconvenienced. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development: Vol. 2. Proceedings from the 1993 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 251–264). Dayton, OH: Ohio Psychology Press.

Lohman, D. F. (1996). Spatial ability and G. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and assessment* (pp. 97–116). Hillsdale, NJ: Erlbaum.

Lohman, D. F. (2005). The role of nonverbal ability tests in identifying academically gifted students: An aptitude perspective. *Gifted Child Quarterly, 49,* 111–138.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points." *Annual Review of Psychology, 51,* 405–444.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904). "'General intelligence,' objectively determined and measured." *Journal of Personality and Social Psychology, 86,* 96–111.

Lubinski, D., & Benbow, C. P. (2000). States of excellence. *American Psychologist, 55,* 137–150.

Lubinski, D., & Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35 years: Uncovering antecedents for the development of math–science expertise. *Perspectives on Psychological Science, 1,* 316–345.

Lubinski, D., & Dawis, R. V. (1992). Aptitudes, skills, and proficiencies. In M. D. Dunnette & L. M. Hough (Eds.), *The handbook of industrial/organizational psychology* (2nd ed., pp. 1–59). Palo Alto, CA: Consulting Psychologists Press.

Lubinski, D., & Humphreys, L. G. (1990a). A broadly based analysis of mathematical giftedness. *Intelligence, 14,* 327–355.

Lubinski, D., & Humphreys, L. G. (1990b). Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial visualization and mathematical ability. *Psychological Bulletin, 107,* 385–393.

Lubinski, D., & Humphreys, L. G. (1996). Seeing the forest from the trees: When predicting the behavior or status of groups, correlate means. *Psychology, Public Policy, and Law, 2,* 363–376.

Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence, 24,* 159–201.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70,* 151–159.

Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Cicchetti & W. Grove (Eds.), *Thinking clearly about psychology* (pp. 3–39). Minneapolis: University of Minnesota Press.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry, 1,* 108–141.

National Academy of Sciences. (2005). *Rising above the gathering storm.* Washington, DC: National Academy Press.

Newcombe, N. S., & Uttal, D. H. (2006). Whorf versus Socrates, round 10. *Trends in Cognitive Sciences, 10,* 394–396.

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science, 10,* 464–480.

Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting intellectual patterns for creativity in the arts and sciences: Tracking intellectually precocious youth over 25 years. *Psychological Science, 18,* 948–952.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science, 19,* 957–961.

Paterson, D. G. (1957). The conservation of human talent. *American Psychologist, 12,* 134–144.

Paterson, D. G., Elliott, R. M., Anderson, L. D., & Toops, H. A. (1930). *Minnesota Mechanical Ability Tests.* Minneapolis: University of Minnesota Press.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37,* 239–253.

Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology, 85,* 112–118.

Schmidt, D. B., Lubinski, D., & Benbow, C. P. (1998). Validity of assessing educational–vocational preference dimensions among intellectually talented 13-year-olds. *Journal of Counseling Psychology, 45,* 436–453.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262–274.

Shaycoft, M. F. (1967). *The high school years: Growth in cognitive skills.* Pittsburgh, PA: School of Education, University of Pittsburgh, American Institutes for Research.

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93,* 604–614.

Silverman, I., Choi, J., & Peters, M. (2007). The hunter–gatherer theory of sex differences in spatial abilities: Data from 40 countries. *Archives of Sexual Behavior, 36,* 261–268.

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

Snow, R. E. (1999). Commentary: Expanding the breadth and depth of admissions testing. In S. Messick (Ed.), *Assessment in higher education* (pp. 133–140). Hillsdale, NJ: Erlbaum.

Snow, R. E., Corno, L., & Jackson, D. N., III. (1996). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 243–310). New York: MacMillan.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263–332). New York: Macmillan.

Stanley, J. C. (1996). In the beginning: The Study of Mathematically Precocious Youth. In C. P. Benbow & D. Lubinski (Eds.), *Intellectual talent* (pp. 225–235). Baltimore: Johns Hopkins University Press.

Stanley, J. C. (2000). Helping students learn only what they don't already know. *Psychology, Public Policy, and Law, 6,* 216–222.

Steen, L. A. (1988, April 29). The science of patterns. *Science, 240,* 611–616.

Super, D. E., & Bachrach, P. B. (1957). *Scientific careers and vocational development theory.* New York: Bureau of Publications, Teachers College, Columbia University.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47,* 599–604.

Vernon, P. A. (Ed.). (1961). *The structure of human abilities* (2nd ed.). London: Methuen London.

Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youth: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology, 94,* 785–794.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology, 99,* 397–420.

Wise, L. L., McLaughlin, D. H., & Steel, L. (1979). *The Project TALENT data bank.* Palo Alto, CA: American Institutes for Research.

(*Appendixes follow*)

Appendix A

Breakdown by Sex of the Degrees and Occupations Included in Figure 5 and Appendix B

Degrees	Occupations		

Degrees

Engineering
  Engineering (1137/336/71, 6/3/0)
Physical science
  Biochemistry (8/9/21, 8/8/5)

  Chemistry (244/43/100, 95/10/9)
  Physical science (122/37/9, 21/5/0)
  Physics (131/69/58, 4/1/0)

Math/computer science
  Computer science (27/42/6, 2/9/0)
  Math (474/128/42, 365/74/3)

  Statistics (7/11/6, 2/2/0)
Biological science
  Anatomy (12/11/16, 8/6/3)
  Biological science (332/69/28, 209/39/6)

  Botany (15/15/7, 7/11/3)

  Zoology (112/20/14, 45/11/2)

Humanities

  English (344/110/24, 1056/140/6)

  Foreign language (111/37/7, 330/84/3)

  History (635/146/24, 374/62/3)

  Humanities (12/3/0, 21/5/1)

  Journalism (89/34/0, 53/8/0)

  Philosophy (110/21/8, 29/1/2)

  Religion (30/39/4, 32/5/0)

Social science

  Economics (410/62/16, 51/3/2)

  Political science (359/72/17, 147/19/3)
  Psychology (338/142/71, 311/66/26)
  Social science (176/43/5, 195/21/4)

  Sociology (234/38/12, 388/18/2)

Visual arts

  Architecture (53/6/0, 6/0/0)

  Fine arts (88/29/0, 219/29/0)

Occupations

Engineering
  Engineer (NEC) (219, 1)
  Civil engineer (163, 2)
  Electrical engineer (315, 0)
  Mechanical engineer (192, 1)
  Aeronautical engineer (60, 0)
  Chemical engineer (78, 1)

Physical science
  Physical scientist (NEC) (20, 1)
  Chemist (153, 19)
  Physicist (62, 2)
  Astronomer (1, 0)
  Geologist (38, 3)
  Meteorologist (10, 0)
  Biochemist (26, 12)

Math/computer science
  Mathematician (38, 10)
  Statistician (28, 9)
  Systems analyst (258, 68)
  Computer programmer (175, 61)
  Computer specialist (NEC) (63, 1)

Biological science
  Biological scientist (NEC) (23, 10)
  Pharmacologist (8, 1)
  Microbiologist (19, 16)

Humanities
  Writer (NEC) (15, 8)
  Fiction writer (1, 1)
  Nonfiction writer (16, 20)
  Journalists/reporters (31, 16)
  Radio–TV reporter (6, 6)
  Publisher (4, 2)
  Editor (39, 39)

Visual arts
  Architect (44, 1)
  Painter/sculptor (4, 5)
  Commercial/fashion artist (18, 24)
  Draftsman (13, 1)
  Photographer (15, 3)
  Interior designer (3, 11)
  Landscape architect (10, 0)
Performing arts
  Musician (instrumental) (15, 9)
  Singer (4, 0)
  Music-related (NEC) (5, 6)
  Dancer, choreographer (0, 9)
  Actor (1, 4)
  Theatrical director (7, 1)
  Theater occupation (NEC) (20, 7)
  Performer (NEC) (2, 0)
  Radio or TV announcer (3, 0)
  Performing artist (NEC) (6, 5)
Business
  Executive (NEC) (13, 1)
  In business for self (NEC) (37, 2)
  Industry, business, or commerce (3, 0)
  Real estate (NEC) (9, 0)
  Insurance (NEC) (3, 0)
  Market analyst (29, 8)
  Banking and finance (196, 8)
  Investment consultant (39, 5)
  CPA (282, 7)
  Accountant or auditor (406, 22)
  Purchasing and procurement (96, 8)
  Buyer for retail store (23, 9)
  Efficiency expert (NEC) (242, 3)

Contractor (62, 0)
Franchiser (123, 26)
Salesclerk/cashier (42, 31)
Routeman (5, 0)
Stockbroker (53, 3)
Securities dealer (15, 0)
Real estate salesman (77, 15)
Insurance salesman (137, 1)
Auto salesman (10, 0)
Other salesmen (NEC) (560, 24)
Sales manager (241, 15)
Education
  Teaching (NEC) (269, 352)
  Teaching young children (1, 9)
  Teaching preschool children (2, 165)
  Teaching elementary school (273, 1232)
  Teaching high school (NEC) (173, 173)
  Teaching high school math (195, 103)
  Teaching high school science (236, 74)
  Teaching high school social studies (205, 96)
  Teaching high school English (122, 259)
  Teaching high school foreign language (51, 76)
  Teaching high school commercial education (49, 63)
  Teaching high school home economics (0, 84)
  Teaching high school trade education (141, 11)
  Teaching high school physical education (117, 101)
  Teaching art (46, 67)
  Teaching music (55, 78)
  Teaching speech in high school (8, 9)
  Teaching the handicapped (71, 147)
  Speech therapist (11, 38)
  School administrator (not college) (124, 40)
  Reading specialist (16, 54)

Appendix A (*continued*)

Degrees		Occupations	
Performing arts	Translator/linguist (4, 4)	Advertiser (47, 17)	Other education specialist (46, 39)
Music (63/40/3, 128/28/1)	University teacher: English (53, 37)	Public relations (41, 34)	Teacher's aide (13, 44)
Performing arts (18/18/1, 40/16/0)	University teacher: foreign language (16, 19)	Personnel administrator (233, 53)	Medicine
Business	Social science	Appraiser/estimator (118, 15)	MD general practitioner (34, 0)
Accounting (708/32/4, 36/1/1)	Psychologist (97, 39)	Credit investigator (137, 15)	MD surgeon (56, 2)
Business and commerce (1393/134/9, 249/9/1)	Economist (17, 3)	Manager and administrator (NEC) (515, 50)	MD psychiatrist (38, 11)
Education	Sociologist (1, 0)	Manufacturing manager (56, 1)	MD medical researcher (12, 6)
Education other (232/441/21, 378/425/2)	Social scientist (NEC) (40, 22)	Retail trade manager (169, 18)	MD other and unspecified (338, 21)
Elementary education (164/80/1, 2149/414/0)	University teacher: social science (109, 29)	Private business agent (3, 0)	Dentistry
Physical education (258/66/2, 222/51/2)		Developer (14, 0)	Dentist (189, 1)
Specific doctorates		Business supervisor (117, 6)	Law
JD (939, 42); DDS (214, 1); MD (490, 48); EdD (43, 17)			Lawyer (777, 28)

Note. This table includes the breakdown by sex of the degrees and occupations included in Figure 5 and Appendix A. In the degrees column, the respective sample sizes are given for bachelors, masters, and doctorates for males and females, respectively (B/M/D for males, B/M/D for females). In the occupations columns, the sample sizes (males/females) are reported. The specific doctorates category in the degrees column and the occupations columns pertain only to Figure 5. Figure B1 includes data from the remainder of the degrees column (i.e., Engineering through Education). NEC = not elsewhere classified.

Appendix B

Average *Z* Scores of Participants on Both General Ability Level and Spatial, Mathematical, and Verbal Ability Level for Bachelor's Degrees, Master's Degrees, and PhD Degrees Plotted by Field

It is important to note the importance of spatial ability for those securing degrees in math/computer science, physical science, and engineering. Hegarty and Waller (2005, p. 155) discussed the importance of spatial ability in the performance of surgeons. Bingham (1937) anticipated this topic, noting that, for surgeons and dentists,

> quite as indispensable is aptitude for visualizing vividly in three dimensions; for it is necessary to see in their true positions and to manipulate the forms observed in a dentist's little mirror or in a laryngoscope; also to picture correctly the highly complicated unseen structures beneath the body surface—arteries, nerves, muscles, tendons, joints, glands, vital organs—perhaps at the end of a probe. (p. 172)

We conducted an analysis using the surgeons and other MDs that can be found here and in Appendix A (MD surgeon, $n = 58$; MD all others, $n = 460$). The difference between the surgeons (avg. $z = 1.17$) and the remainder (avg. $z = 1.12$) on spatial ability was 0.05. The highest in spatial ability were the MD medical researchers (avg. $z = 1.27$) in comparison to all other subgroups. Spatial ability is evidently important not only for surgeons but all the medical fields examined in Project TALENT, and in particular for medical research.

We conducted an analysis to determine the similarity between all four cohorts compared to the 9th-grade cohort alone. The average difference across all four grades combined and the 9th-grade sample was less than the absolute value of 0.08. For completeness, the *g* level (average of S + M + V) of the bachelors (BA and BS) and doctorates (PhDs) was computed within each group. In the order corresponding to the graph, these were as follows: engineering (PhD = 1.73; bachelors = 1.22), physical science (PhD = 1.62; bachelors = 1.15), math/computer science (PhD = 1.75; bachelors = 1.18), biological science (PhD = 1.33; bachelors = 0.86), humanities (PhD = 1.34; bachelors = 0.84), social science (PhD = 1.29; bachelors = 0.75), arts (masters + PhD = 0.97; bachelors = 0.71), business (masters + PhD = 0.99; bachelors = 0.64), and education (masters + PhD = 0.64; bachelors = 0.46).
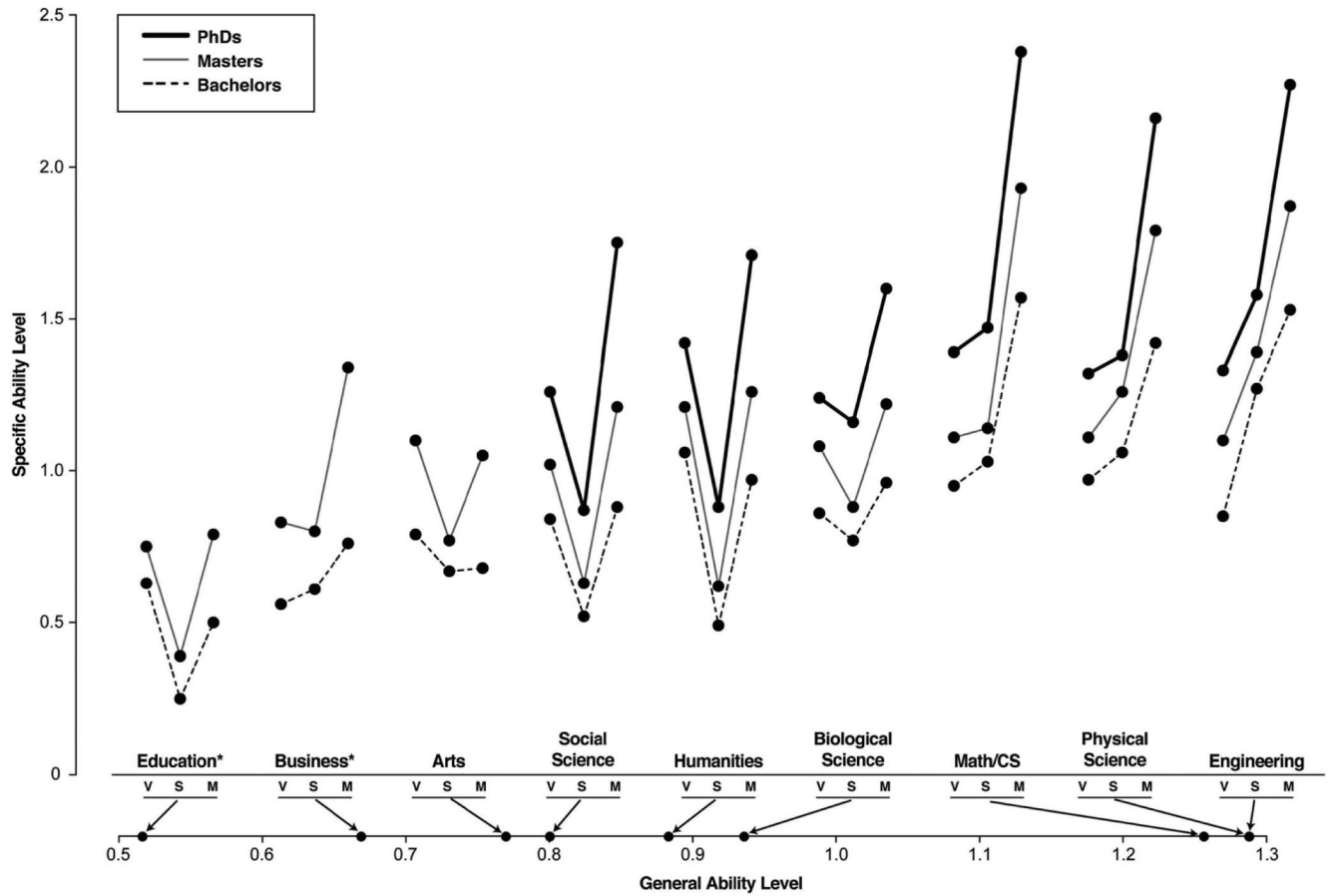
(*Appendixes continue*)

Figure B1.   * For education and business, masters and doctorates were combined because the doctorate samples for these groups were too small to obtain stability ($n < 30$). For the specific n for each degree by sex that composed the major groupings, see Appendix A. Average z scores of participants on spatial, mathematical, and verbal ability for bachelor's degrees, master's degrees, and PhDs are plotted by field in Figure B1. The groups are plotted in rank order of their normative standing on g (verbal [V] + spatial [S] + mathematical [M]) along the x-axis, and each arrow indicates on the continuous scale where each field lies on general mental ability. All x-axis values are based on the weighted means across each degree grouping. This figure is standardized in relation to all participants with complete ability data at the time of initial testing. Respective ns for each group (males + females) were as follows (for bachelor's, master's, and doctorates, respectively): engineering (1,143, 339, 71), physical science (633, 182, 202), math/computer science (877, 266, 57), biological science (740, 182, 79), humanities (3,226, 695, 82), social science (2,609, 484, 158), arts (615, masters + doctorates = 171), business (2,386, masters + doctorates = 191), and education (3,403, masters + doctorates = 1,505).

Appendix C

Spatial Ability Composite Means and Standard Deviations by Socioeconomic Status (SES) Quartile, Grade, and Sex

SES quartile	9th grade		10th grade		11th grade		12th grade	
	Males	Females	Males	Females	Males	Females	Males	Females
1	60.67 (20.70) 13,056	50.99 (17.66) 12,196	65.16 (21.68) 12,642	54.36 (18.98) 12,784	69.95 (22.16) 11,502	57.57 (19.52) 12,101	72.79 (22.73) 9,263	60.05 (20.10) 9,900
2	68.06 (21.10) 11,514	58.68 (18.54) 11,509	73.77 (21.19) 11,890	62.57 (19.66) 11,825	78.68 (21.63) 11,095	65.22 (19.31) 10,167	81.90 (21.55) 9,504	67.75 (19.83) 10,609
3	72.75 (21.31) 11,512	63.43 (18.77) 11,781	78.33 (21.11) 11,477	67.21 (19.38) 10,046	83.03 (20.86) 9,187	69.55 (19.54) 11,581	86.30 (21.02) 9,493	71.82 (19.73) 8,705
4	78.80 (20.92) 11,123	69.23 (18.98) 11,883	83.97 (20.69) 10,696	72.20 (19.77) 11,147	87.59 (20.61) 10,368	74.98 (19.78) 10,256	91.52 (20.33) 8,430	76.57 (20.12) 9,432

Note. In each cell, the mean, standard deviation (in parentheses) and *n* are reported.