

## Research Report

# Contrasting Intellectual Patterns Predict Creativity in the Arts and Sciences

## Tracking Intellectually Precocious Youth Over 25 Years

Gregory Park, David Lubinski, and Camilla P. Benbow

Vanderbilt University

---

**ABSTRACT**—A sample of 2,409 intellectually talented adolescents (top 1%) who were assessed on the SAT by age 13 was tracked longitudinally for more than 25 years. Their creative accomplishments, with particular emphasis on literary achievement and scientific-technical innovation, were examined as a function of ability level (sum of math and verbal SAT scores) and tilt (math SAT score minus verbal SAT score). Results showed that distinct ability patterns uncovered by age 13 portend contrasting forms of creative expression by middle age. Whereas ability level contributes significantly to creative accomplishments, ability tilt is critical for predicting the specific domain in which they occur (e.g., securing a tenure-track position in the humanities vs. science, technology, engineering, or mathematics; publishing a novel vs. securing a patent).

---

Modern global economies are now knowledge based and driven largely by intangibles such as creativity, information, and ideas, rather than natural resources or physical capital (Florida, 2002; Friedman, 2005; Stewart, 2001; Suarez-Villa, 2000). Multiple disciplines and human-capital initiatives have noted that the rapidly changing economy is led by the *creative sector*, a diverse group of occupations including, but not limited to, work in science, technology, the arts, medicine, law, and entertainment. This sector is projected to generate approximately 10 million new jobs between 2004 and 2014 (Florida, 2006). The increased

mobility of individuals working in this sector, combined with the increased speed of information transfer between employers and potential employees, has globalized competition for innovative talent, and several disciplines and industries are investing in new methods to identify and secure exceptional human capital (American Competitiveness Initiative, 2006; Friedman, 2005). However, few disciplines are taking into account the dimensions of human individuality that give rise to creative expression. The current study examined the hypothesis that contrasting patterns of intellectual precocity manifest in early adolescence engender qualitatively different forms of creativity by middle age. Two broad domains were examined: the humanities versus science, technology, engineering, and mathematics (STEM).

Recently, empirical findings have shown that individual differences within the top 1% of ability predict differences in occupational performance and creativity: More ability increases the likelihood of accomplishments such as earning a doctorate, earning tenure at a top-50 U.S. university, earning a high income, and securing a patent (Lubinski, Benbow, Webb, & Bleske-Rechek, 2006; Wai, Lubinski, & Benbow, 2005). Most normative assessments, however, are unable to differentiate the able from the exceptionally able, because both groups tend to pile up at the ceiling of conventional indicators such as college entrance exams. The lack of variation at the upper end constrains the covariation between these measures and subsequent accomplishments. When college entrance exams are administered to the intellectually precocious before age 13, however, these youth generate score distributions like those of typical college-going 12th graders, and the able and exceptionally able are readily distinguished (Lubinski & Benbow, 2006). When these youth are tracked over multiple decades, the psychological import of individual differences within the top 1%, which spans

---

Address correspondence to Gregory Park, David Lubinski, or Camilla P. Benbow, Department of Psychology and Human Development, Peabody 512, Vanderbilt University, Nashville, TN 37203, e-mail: greg.park@vanderbilt.edu, david.lubinski@vanderbilt.edu, or camilla.benbow@vanderbilt.edu.

more than one third of the ability range, becomes open to evaluation. For example, IQs in the top 1% begin at approximately 137 and extend beyond 200. But in this case, too, outcome criteria with high ceilings are required to appraise the validity of these early assessments longitudinally (and follow-up intervals must be sufficiently long to allow for the development of the expertise needed for creative accomplishments).

In the study reported here, we tested the hypothesis that among intellectually precocious youth within the top 1% of ability, the pattern of exceptional mathematical and verbal reasoning abilities, as assessed at age 12, differentially predict creative achievements in the humanities versus STEM domains 25 years later.

## PARTICIPANTS

Participants were drawn from 20-year follow-ups of the first three cohorts of the Study of Mathematically Precocious Youth's (SMPY's) planned 50-year longitudinal study of intellectual talent (Lubinski & Benbow, 2006). Through talent searches, youth took the SAT before age 13, and those who were in the top 1% of ability for their age were selected for participation in the SMPY study. (The combined sample included 1,569 males and 840 females.)

Twenty years after their identification and initial assessment, participants were surveyed (at approximately age 33) through mailed questionnaires, phone interviews, or Internet surveys. The collection of these 20-year data occurred between 1992 and 1994 for Cohort 1 (Benbow, Lubinski, Shea, & Eftekhari-Sanjani, 2000), between 1996 and 1999 for Cohort 2 (Benbow et al., 2000), and in 2003 and 2004 for Cohort 3 (Lubinski et al., 2006). The 20-year follow-up surveys included questions about educational and occupational achievements, as well as family and lifestyle. Response rates ranged between 77% and 82% across cohorts.

For participants who reported professorial positions or who had secured doctorates by the time of their 20-year follow-up, we ascertained professional status for the 2005–2006 academic year through university Web sites. We used U.S. News & World Report's (2006) listing of *America's Best Colleges* to generate a reasonable list of the top 50 U.S. universities. To update the achievements reported in the 20-year follow-up surveys, we used Internet databases to collect current data on patents and literary achievements. Patent data were secured using Google patents ([www.google.com/patents](http://www.google.com/patents)), and information on literary publications was secured through Amazon ([www.amazon.com](http://www.amazon.com)). We limited literary achievements to published novels, collections of short stories, regular columns in current periodicals, nonfiction books (not including technical or instructional guides), and produced screenplays and dramatic plays. This part of the follow-up took place at least 25 years after the participants' initial identification.

## DESIGN

Math and verbal SAT scores secured by age 13 were transformed into two relatively independent dimensions ( $r = .02$ ), which were subsequently transformed into  $z$  scores: ability level (sum of the math and verbal scores) and ability tilt (math score minus verbal score). The former assessed general ability level; the latter, differential ability strength. Positive ability tilt indicated greater strength in quantitative than verbal ability, whereas negative ability tilt reflected stronger verbal than quantitative ability.

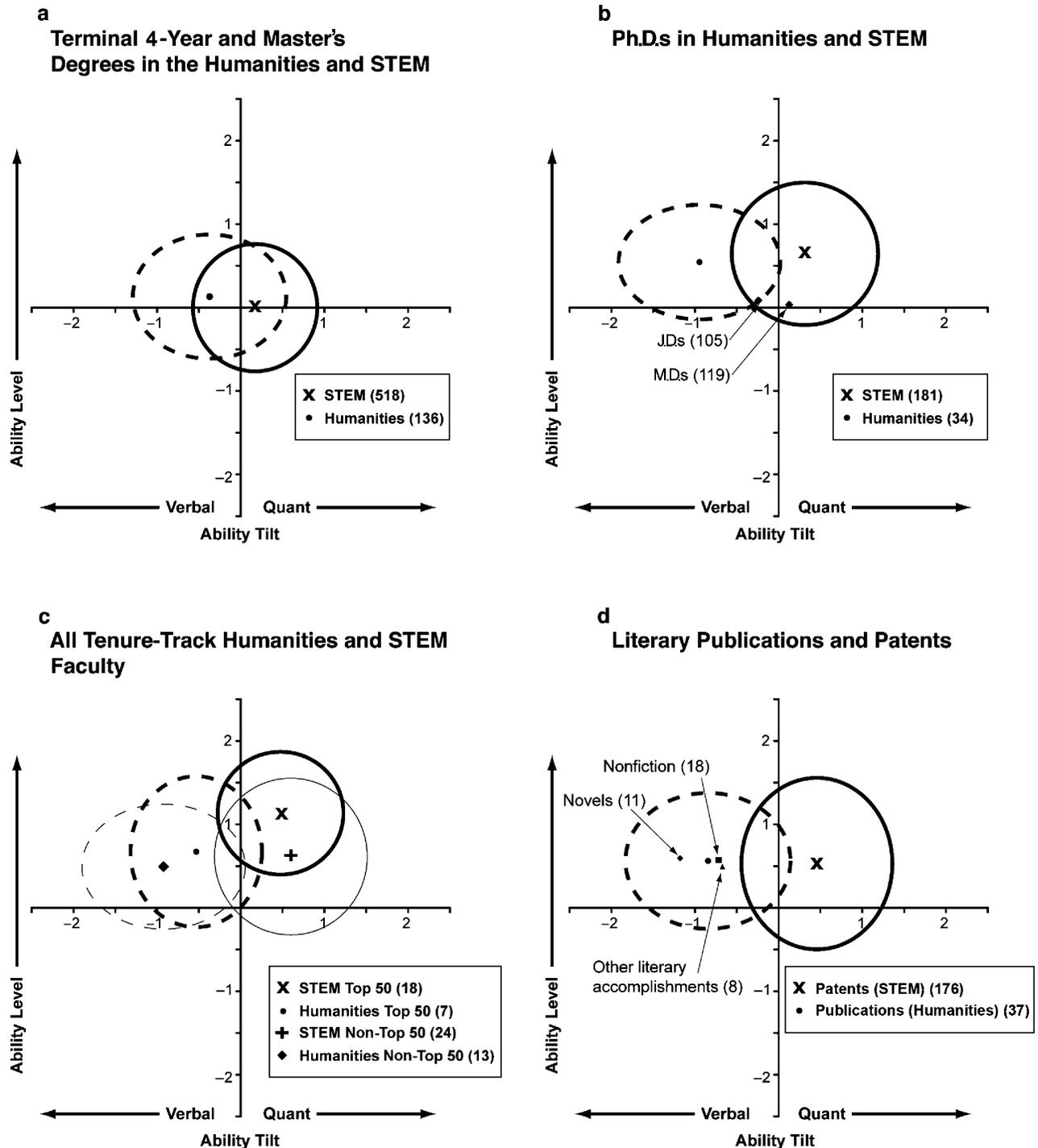
Participants were classified according to four broad categories of accomplishment: securing a terminal bachelor's or master's degree (Fig. 1a), securing a doctorate (Ph.D.; Fig. 1b), securing a tenure-track position at a U.S. university (Fig. 1c), and securing a patent or authoring a noteworthy literary publication (Fig. 1d). For each category, we distinguished achievements in STEM versus the humanities. Patents and publications were classified as STEM and humanities accomplishments, respectively. STEM degrees included the physical sciences, mathematics, computer science, and engineering. Humanities degrees included art, history, literature, languages, drama, and related fields. Tenure-track positions in these fields were classified accordingly as STEM or humanities positions. (Other fields, such as the social sciences, biological sciences, health sciences, architecture, business, and management, were not analyzed for the purposes of this study.)

## RESULTS

Each panel in Figure 1 represents the two-dimensional space defined by ability tilt ( $x$ -axis) and ability level ( $y$ -axis). Bivariate means for the humanities and STEM groups are plotted, and the ellipses represent the space within 1 standard deviation of the means on each dimension. Figure 1c shows four, rather than two, ellipses in order to distinguish participants who secured tenure-track positions at top-50 U.S. universities from those with tenure-track positions at other U.S. universities. Figures 1b and 1d show bivariate means, without ellipses, for additional specific criterion groups (participants who secured a J.D. or M.D.; novelists and nonfiction authors), in order to provide a more complete portrait of the accomplishments of this sample.

Examination of Figure 1 confirms that the humanities and STEM groups occupy different regions in the space defined by ability tilt and ability level. Like most powerful findings, these are readily seen by the naked eye. Even so, we performed statistical analyses to quantify the degree of separation between the humanities and STEM groups and to test for significance.

Within each broad achievement category, we contrasted the STEM group with the humanities group, using the  $d$  statistic (Cohen, 1988) to measure the magnitude of their difference in ability tilt and ability level. No statistically significant differences were observed for ability level. For ability tilt, however,  $t$  tests indicated that the STEM and humanities groups differed significantly ( $p_{\text{rep}} > .99$ ) in every comparison: terminal bachelor's



**Fig. 1.** Participants' achievements as a function of ability tilt (math SAT score minus verbal SAT score) and ability level (sum of the math and verbal SAT scores), in standard deviation units. The achievement categories examined were (a) completing a terminal 4-year or master's degree, (b) completing a Ph.D. (means for M.D.s and J.D.s are also shown), (c) securing a tenure-track faculty position, and (d) publishing a literary work or securing a patent. In each graph, bivariate means are shown for achievements in humanities and in science, technology, engineering, and mathematics (STEM), respectively; the ellipse surrounding each mean indicates the space within 1 standard deviation on each dimension. The  $n$  for each group is indicated in parentheses. The mean SAT scores (math and then verbal) for the criterion groups were as follows: 4-year and master's STEM degree—575, 450; 4-year and master's humanities degree—551, 497; STEM Ph.D.—642, 499; humanities Ph.D.—553, 572; tenure-track STEM position in a top-50 university—697, 534; tenure-track humanities position in a top-50 university—591, 557; tenure-track STEM position in a non-top-50 university—659, 478; tenure-track humanities position in a non-top-50 university—550, 566; patents (i.e., STEM creative achievements)—626, 471; and publications (i.e., humanities creative achievements)—561, 567.

and master's degrees ( $d = 0.71$ ), Ph.D.s ( $d = 0.63$ ), tenure-track positions ( $d = 1.62$ ), and patents and published novels and literary publications ( $d = 1.67$ ). The contrast for tenure-track positions combined top-50 and lower-ranked schools, but notice how the ellipses for the top schools converge in Figure 1c. This convergence is due to a number of participants who had earned very high scores on the math portion of the SAT. For example, the mean math SAT score of the 18 participants who later earned tenure-track positions in STEM fields at top-50 U.S. universities was 697, and the lowest score in this group was 580 (a score greater than that of more than 60% of all participants); these high scores suggest that the tilt for this group was reduced because of the ceiling on SAT math scores. Two individuals earned the top possible score (800), which illustrates that for profoundly gifted participants, college entrance exams such as the SAT can manifest ceiling effects as early as age 12 (cf. Benbow & Stanley, 1996; Muratori et al., 2006; Stanley, 2000).

To quantify the distinctiveness of each ellipse in Figure 1, we conducted contrasts comparing each criterion group with the remainder of the sample on both ability level and ability tilt. For example, we compared the 34 participants who earned Ph.D.s in the humanities with the remaining 2,375 participants, and repeated this procedure for all other criterion groups. Statistical significance was evaluated by  $t$  tests, and  $d$  was computed to measure the size of the difference between each group and the remainder of the sample. Every humanities and STEM group was significantly different ( $p_{\text{rep}} > .98$ ) from the remainder of the sample on both ability level and tilt, with the exception of the groups defined by terminal 4-year and master's degrees; neither the group with STEM degrees nor the group with humanities degrees was significantly different from the remainder of the sample on ability level. The significant effect sizes for the differences between the criterion groups and the remainder of the sample ranged from  $-1.13$  (literary publications) to  $0.69$  (STEM tenure-track positions) for ability tilt and from  $0.68$  (humanities Ph.D.s) to  $1.09$  (STEM tenure-track positions) for ability level.

Sex differences were also observed, as the males had a greater quantitative tilt than the females ( $d = 0.72$ ,  $p_{\text{rep}} > .99$ ), and males scored higher than females on ability level ( $d = 0.40$ ,  $p_{\text{rep}} > .99$ ). These differences are reflected in the percentage of males and females in each criterion group: terminal 4-year and master's degrees in STEM (males: 24.0%, females: 16.9%), terminal 4-year and master's degrees in humanities (males: 4.4%, females: 8.0%), STEM Ph.D.s (males: 9.9%, females: 3.0%), humanities Ph.D.s (males: 1.1%, females: 1.9%), M.D.s (males: 4.5%, females: 5.8%), J.D.s (males: 4.3%, females: 4.5%), tenure-track STEM positions (males: 2.5%, females: 0.4%), tenure-track humanities positions (males: 0.6%, females: 1.3%), literary publications (males: 1.3%, females: 2.0%), and patents (males: 10.6%, females: 1.3%).

Finally, the creative potential of these talent-search participants is further underscored by the fact that, overall, this sample earned a total of 817 patents and published 93 books (56 novels,

37 nonfiction books). Last year, one participant was awarded the Fields Medal (thought of as the Nobel Prize for mathematics), and this year, another participant won the John Bates Clark Medal (most outstanding economist under 40).

## DISCUSSION

Distinct ability patterns among intellectually precocious youth foreshadow creative accomplishments by middle age. Although ability level is informative for predicting overall achievement and creativity in general terms (Benbow, 1992; Lubinski et al., 2006; Wai et al., 2005), ability tilt contributes to the prediction of the domain in which exceptional accomplishments are likely to occur.

That a 3-hr assessment conducted by age 13 captures individual differences that make a difference in forecasting rare accomplishments, creative achievements, and qualitatively different developmental trajectories is important for many reasons. Yet several recent statements in highly visible outlets have asserted that there is little evidence that high scores on standardized instruments, such as the SAT, relate to real-world success later in life, particularly in science and technology careers:

- “There is little evidence that those scoring at the very top of the range in standardized tests are likely to have more successful careers in the sciences” (Muller et al., 2005, p. 1043).
- “Measures of aptitude for high school and college science have not proved to be predictive of success in later science and engineering careers” (Committee on Maximizing the Potential of Women in Academic Science and Engineering, and Institute of Medicine, 2007, p. 25).
- “Standardized tests are thus not sufficiently predictive of future performance. Individuals are not necessarily more meritorious if they obtain the highest scores on standardized tests, thus rendering invalid the argument that students with the highest scores should have priority in admissions” (Vasquez & Jones, 2006, p. 138).

Our results falsify these statements (also see Friedman, 2005, pp. 266–267).

Even more refined predictions of contrasting forms of creativity would likely result from including spatial ability as a predictor in addition to quantitative and verbal ability. Other longitudinal studies have revealed that all three of these abilities possess incremental validity relative to the other two in identifying psychologically significant forms of intellectual talent and in forecasting remote educational and occupational accomplishments (Gohm, Humphreys, & Yao, 1998; Humphreys, Lubinski, & Yao, 1993; Shea, Lubinski, & Benbow, 2001; Webb, Lubinski, & Benbow, 2007).

Exceptionally high scores on standardized measures of cognitive abilities are informative and highly significant psychologically. However, the individual differences in this study were uncovered by administering college entrance exams to 12-year-olds; such individual differences are routinely veiled

when age-appropriate assessments are administered to intellectually precocious youth, such as when they take the SAT in high school. When intellectually talented students reach this stage of development, essentially all of their scores cluster near the ceiling—and the exceptionally able are no longer readily distinguished from the able. Thus, all too often, first-rate engineering and physical science faculty indicate that math scores on the SAT “don’t mean much,” because all of their applicants score in the top 700s.<sup>1</sup> When artificial ceilings are imposed on psychometric measures (or physical measures), variation is constrained, and, therefore, the covariation between such measures and meaningful criteria is severely limited.

To adequately reveal the psychological significance of individual differences within the top 1% of ability (which covers more than one third of the total ability range), and to empirically validate the assessment tools that measure these individual differences, research must use large sample sizes, incorporate measures with high ceilings, adopt criteria with high ceilings or low base rates, and allow enough time for creative achievements to develop. When these design features are in place, it becomes possible to appraise differential capabilities in intellectually talented populations, and the creative promise that these individual differences harbor is revealed.

In conclusion, challenging ability tests administered by age 13 to highly able students can predict their creative production 25 years later and the specific nature of these accomplishments.

**Acknowledgments**—This article is based on a thesis submitted by Gregory Park to Vanderbilt University in partial fulfillment of the master’s of science degree. An earlier version of this article profited from comments by Kimberley Ferriman, Jonathan Wai, and Leslie J. Yonce. Support was provided by a Research and Training Grant from the Templeton Foundation and by National Institute of Child Health and Human Development Grant P30 HD 15052 to the Vanderbilt Kennedy Center for Research on Human Development.

## REFERENCES

- 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. (1992). Academic achievement in math and science between ages 13 and 23: Are there differences in the top one percent of ability? *Journal of Educational Psychology, 84*, 51–61.
- Benbow, C.P., Lubinski, D., Shea, D.L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability: Their status 20 years later. *Psychological Science, 11*, 474–480.
- 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.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Committee on Maximizing the Potential of Women in Academic Science and Engineering, and Institute of Medicine. (2007). *Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering*. Washington, DC: National Academies Press.
- Florida, R. (2002). *The rise of the creative class*. New York: Basic Books.
- Florida, R. (2006, September 15). Regions and universities together can foster a creative economy. *Chronicle of Higher Education*, p. B6.
- Friedman, T.L. (2005). *The world is flat*. New York: Farrar, Straus, & Giroux.
- Gohm, C.L., Humphreys, L.G., & Yao, G. (1998). Underachievement among spatially gifted students. *American Educational Research Journal, 35*, 515–531.
- 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.
- 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., Benbow, C.P., Webb, R.M., & Bleske-Rechek, A. (2006). Tracking exceptional human capital over two decades. *Psychological Science, 17*, 194–199.
- Muller, C.B., Ride, S.M., Fouke, J., Whitney, T., Denton, D.D., Cantor, N., et al. (2005). Gender differences and performance in science. *Science, 307*, 1043.
- Muratori, M.C., Stanley, J.C., Gross, M.U.M., Ng, L., Tao, T., Ng, J., & Tao, B. (2006). Insights from SMPY’s greatest former prodigies: Drs. Terence (“Terry”) Tao and Lenhard (“Lenny”) Ng reflect on their talent development. *Gifted Child Quarterly, 50*, 307–324.
- 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.
- Stanley, J.C. (2000). Helping students learn only what they don’t already know. *Psychology, Public Policy, and Law, 6*, 216–222.
- Stewart, T.A. (2001). *The wealth of knowledge: Intellectual capital and the twenty-first century organization*. New York: Basic Books.
- Suarez-Villa, L. (2000). *Invention and the rise of technocapitalism*. Lanham, MD: Rowman & Littlefield.
- U.S. News & World Report. (2006). *America’s best colleges 2007*. Retrieved September 2006 from [http://www.usnews.com/usnews/edu/college/rankings/brief/t1natudoc\\_brief.php](http://www.usnews.com/usnews/edu/college/rankings/brief/t1natudoc_brief.php)
- Vasquez, M.J.T., & Jones, J.M. (2006). Increasing the number of psychologists of color: Public policy issues for affirmative diversity. *American Psychologist, 61*, 132–142.
- 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, 97*, 484–492.
- 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.

<sup>1</sup>Of course, high school SAT scores in the high 700s are meaningful in the sense that students who earn such scores possess much more potential than, say, students clustering around the high 500s. However, to distinguish the exceptionally able from the able among students who score in the high 700s, one must use measures with higher ceilings or administer assessments like the SAT at an earlier age.