Cognitive Profiles of Verbally and Mathematically Precocious Students: Implications for Identification of the Gifted

Camilla Persson Benbow
Study of Mathematically Precocious Youth (SMPY)
and
Department of Psychology
Iowa State University

Lola L. Minor
Johns Hopkins University

Abstract

Performance on tests of specific abilities commonly associated with intelligence was contrasted between 13-year-olds identified as extremely precocious (top 1 in 10,000) in either verbal or mathematical reasoning ability. Such students differ cognitively. Verbally precocious students scored higher on verbal and general knowledge types of tests, and mathematically precocious students scored higher on tests of nonverbal reasoning, spatial ability, and memory. Results from the factor analysis of test scores (excluding memory test scores) yielded three factors: spatial/speed, verbal, and nonverbal. Mathematically talented students had higher scores on the nonverbal and speed factors; verbally talented students had higher scores on the verbal factor. Thus, at least two distinct forms of giftedness seem to exist (i.e., verbal and nonverbal). Their evolution, moreover, appeared to follow different developmental paths, consistent with Gagne (1985).

Results from a national survey of school personnel prompted the conclusion that “a labyrinth of confusion” concerning the “giftedness” construct is widespread (Richert, Alvino, & McDonnel, 1982). Moreover, Hoge (1988) concluded that serious deficiencies exist in current definitions of “giftedness.” Educators often view giftedness as something requiring a label (Guskin et al., 1986). To further our understanding of “giftedness,” this paper describes the cognitive profiles associated with two types of “giftedness.” Implications of the findings for identification of gifted students is then addressed.

Considerable effort has already been made to develop and refine the giftedness construct (e.g., Dark & Benbow, in press; Feldhusen, 1986; Gagne, 1985; Hagan, 1980; Sternberg, 1981, 1986; Sternberg & Davidson, 1986). Most of this recent work, including some of our own work (Dark & Benbow, in press), has been in the information-processing domain. Yet giftedness, as originally operationalized, emerged out of the psychometric tradition of research (e.g., the work of Terman, Hollingworth, or Guilford). Classifying students as gifted on the basis of cognitive ability test scores is a psychometric procedure.

Putting The Research To Use

The construct of giftedness was investigated. Results indicated that two types of giftedness, verbal and nonverbal, are distinct from one another. Thus, procedures for identifying gifted students should include assessments of both types of talents. Several investigators (e.g., Feldhusen, 1989) have reported that selecting students for a “gifted” program on the basis of one overall ability is indeed questionable. Our results indicated that reliance on global indicators of intellectual functioning may exclude too many nonverbally gifted students, who appear to be less balanced than verbally gifted students in their cognitive development. Gifted students should be selected, therefore, for special programs on the basis of having qualities that match the objectives of the program. Conversely, programs should be developed to serve the educational needs of children with either of these types of giftedness.
rently, the "state of the art of identification of gifted and talent-
ed youth is in some disarray" (Richert, Alvino, & McDonnel,
1982). Standardized achievement tests and tests of general
intelligence are widely used for identification (Yarborough &
Johnson, 1983). Yet there is prevailing skepticism that results
from either form of assessment adequately reflect giftedness
(Feldhusen, 1989). The validity of one overall indicator of
intellectual functioning (e.g., the IQ score) has been ques-
tioned (e.g., Feldhusen, 1989; Feldhusen, Asher, & Hoover,
1984; Gagne, 1985; Gardner, 1983; Parke, 1989; Renzulli,
1984; Stanley, 1984b; Sternberg & Davidson, 1986). In its
stead, a multiple-talent approach has been offered (e.g.,
Donlon, 1984), is an especially good measure of reasoning
among meaningless figures and to develop a systematic
method of reasoning. Designed to measure "clear thinking,"
Raven’s test is often used for cross-cultural testing (i.e., test-
ing persons with highly dissimilar backgrounds) (Anastasi,
1982).

Spatial Ability (Thurstone’s Space) was measured by two
standardized tests which were designed for adolescents and
young adults (Guilford & Zimmerman, 1981) and one ex-
perimental test. The Guilford-Zimmerman Spatial Orientation
test measures the ability to perceive arrangements of items
of visual information in space. The Guilford-Zimmerman Spa-
tial Visualization test requires the cognition of visual transfor-
mations. The transformations are changes in location or
position, rearrangements of parts, or substitutions of one visual
object for another. Cubes, the experimental test, measures
an individual’s ability to form and manipulate mental images
of objects (Benbow et al., 1983). All three spatial ability tests
were highly speeded.

Nonverbal Reasoning (Thurstone’s General Reasoning) was
measured by Raven’s Progressive Matrices, Advanced Set
(Raven, Court, & Raven, 1977). This 36-item, untimed test
measures a person’s capacity to apprehend relationships
among meaningless figures and to develop a systematic
method of reasoning. Designed to measure “clear thinking,”
Raven’s test is often used for cross-cultural testing (i.e., test-
ing persons with highly dissimilar backgrounds) (Anastasi,
1982).

Mechanical Comprehension was measured by the Bennett
Mechanical Comprehension Test, form AA, which was
designed to measure the ability of an individual to understand
various kinds of physical and mechanical relationships (Ben-
nett, 1940).

Vocabulary and General Information Knowledge (Thur-
stone’s Verbal Comprehension) was measured by Terman’s
Concept Mastery Test, Form T. It was designed to test Ter-
man’s group of gifted subjects as adults. It was designed to
measure the ability to deal with abstract ideas at a high level
(Terman, 1956). Some investigators consider it to be a difficult
test of general intelligence and verbal ability.

Methods

Subjects
The Scholastic Aptitude Test (SAT), a test of developed
verbal and mathematical reasoning ability of 17-year-olds
(Donlon, 1984), is an especially good measure of reasoning
among intellectually gifted 12- to 13-year-olds (Stanley &
Benbow, 1986). From November 1980 through October
1983 the Study of Mathematically Precocious Youth (SMPY)
conducted a national talent search for students who scored
at least 700 on SAT-Mathematics before age 13 (Stanley,
1984a). During those three years, 292 such students were
discovered. Almost concurrently, the Center for the Advance-
ment of Academically Talented Youth (CTY) at Johns Hop-
kins conducted a national search for students who scored at
least 630 on SAT-Verbal before age 13. CTY identified 165
students. It was estimated that such students represent the
top 1 in 10,000 of their age group in the respective abilities.
Several students (48) scored at least 630 on SAT-V and 700
on SAT-M before age 13.

As a service for the already identified mathematically pre-
cocious students, three supplemental cognitive testing sessions
were held in May 1981, 1982, and 1983. Another testing
session was held in May 1983 for the verbally precocious stu-
dents identified in CTY’s 1983 Talent Search. Thus, only a
small number of verbally talented students were tested. The
data from these testing sessions were used in this study. A
total of 144 students participated: 106 mathematically talented
(termed 700M’s), 20 verbally talented (termed 630V’s), and
18 who met both the verbal and mathematics criteria (termed
Doubles). At the time of testing subjects were approximately
13-years-old.

Instruments
A battery of tests sufficiently difficult for highly able students
was selected to measure several basic aptitudes. An attempt
was made to measure those primary abilities proposed by
Thurstone (1938) which have been most frequently corrobo-
rated by himself and others. Those primary abilities include:
verbal comprehension, word fluency, number, space, associ-
ative memory, perceptual speed, and general reasoning. We
tested these specific abilities, except numerical ability (since
SAT-M scores were already available) and word fluency
(which seemed unimportant for extremely precocious students
in our sample). We added instead a test of mechanical com-
prehension and a test of language usage, two specific apti-
tudes often included in test batteries. Test reliabilities
approached .9 for the standardization samples and for our
sample.

Spatial Ability (Thurstone’s Space) was measured by two
standardized tests which were designed for adolescents and
young adults (Guilford & Zimmerman, 1981) and one ex-
perimental test. The Guilford-Zimmerman Spatial Orientation
test measures the ability to perceive arrangements of items
of visual information in space. The Guilford-Zimmerman Spa-
tial Visualization test requires the cognition of visual transfor-
mations. The transformations are changes in location or
position, rearrangements of parts, or substitutions of one visual
object for another. Cubes, the experimental test, measures
an individual’s ability to form and manipulate mental images
of objects (Benbow et al., 1983). All three spatial ability tests
were highly speeded.

Nonverbal Reasoning (Thurstone’s General Reasoning) was
measured by Raven’s Progressive Matrices, Advanced Set
(Raven, Court, & Raven, 1977). This 36-item, untimed test
measures a person’s capacity to apprehend relationships
among meaningless figures and to develop a systematic
method of reasoning. Designed to measure “clear thinking,”
Raven’s test is often used for cross-cultural testing (i.e., test-
ing persons with highly dissimilar backgrounds) (Anastasi,
1982).

Mechanical Comprehension was measured by the Bennett
Mechanical Comprehension Test, form AA, which was
designed to measure the ability of an individual to understand
various kinds of physical and mechanical relationships (Ben-
nett, 1940).

Vocabulary and General Information Knowledge (Thur-
stone’s Verbal Comprehension) was measured by Terman’s
Concept Mastery Test, Form T. It was designed to test Ter-
man’s group of gifted subjects as adults. It was designed to
measure the ability to deal with abstract ideas at a high level
(Terman, 1956). Some investigators consider it to be a difficult
test of general intelligence and verbal ability.
Memory (Thurstone's Associative Memory) was measured by using the Coding subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R). It was adapted for group administration with permission from The Psychological Corporation.

Speed (Thurstone's Perceptual Speed) was measured by the Clerical Speed and Accuracy subtest of the Differential Aptitude Test (Form T; Bennett, Seashore, & Wesman, 1974). This 100-item test was designed for students in grades 8 through 12. It tests speed of perception, momentary retention, and speed of response.

Mechanics of English Expression was measured by the Test of Standard Written English (TSWE), which is one of the three parts of the SAT. TSWE has 50 five-option, multiple-choice items to be answered in 30 minutes.

**Procedure**

Data were analyzed by use of the SPSSX computer program. In comparisons between the verbal and mathematical talent groups, the 18 tested students who met the criteria for both groups were excluded. Because of the unequal N's in the subgroups, the ANOVAs in this study were nonorthogonal. It was decided to retain the nonorthogonal design (because the larger the total N, the greater the statistical power) and follow the four-step procedure outlined by Applebaum and Cramer (1974) for nonorthogonal ANOVAs. In addition, the data for the 112 subjects who had no missing scores were combined, submitted to a factor analysis (principal-axis), rotated, and factor scores computed. Effect sizes (Cohen, 1977) were computed for all t-tests. For comparison purposes, Cohen arbitrarily classified effect sizes as being either small ($d \geq 0.2$), medium ($d \geq 0.5$), or large ($d \geq 0.8$).

**Results**

Mean scores of the verbally and mathematically talented students, as well as of those both mathematically and verbally talented, for the various specific aptitude tests are shown by sex in Table 1. The mean scores of the examinees were, for the most part, equivalent to those earned by individuals at least five years older. On the spatial orientation test and especially on the spatial visualization test, these 13-year-olds scored above the average of college students. Even more impressive, however, were the scores on the nonverbal reasoning test. Relative to university students in England, this sample of extremely talented students scored at the 98th percentile.

*In the first testing session (May 1981) a slightly different set of tests was utilized. Thus, the individuals excluded from the factor analysis were almost all 700M males.*

### Table 1

**Performance by Sex of the Three Groups of Extremely Talented Individuals on the Specific Ability Measures**

<table>
<thead>
<tr>
<th></th>
<th>Mathematically Talented Students</th>
<th>Verbally Talented Students</th>
<th>Mathematically and Verbally Talented Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males (N = 13)</td>
<td>Females (N = 7)</td>
<td>Males (N = 10)</td>
</tr>
<tr>
<td></td>
<td>X s.d. (N)</td>
<td>X s.d. (N)</td>
<td>X s.d. (N)</td>
</tr>
<tr>
<td>Guilford-Zimmerman</td>
<td>25.3 9.0 (66)</td>
<td>20.4 9.8 (14)</td>
<td>15.6 7.4</td>
</tr>
<tr>
<td>Orientation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guilford-Zimmerman</td>
<td>24.8 7.7 (66)</td>
<td>20.7 5.0 (14)</td>
<td>18.0 8.2</td>
</tr>
<tr>
<td>Visualization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cubes</td>
<td>24.5 4.0 (91)</td>
<td>23.1 4.6 (15)</td>
<td>15.4 5.2</td>
</tr>
<tr>
<td>Raven's P.M.</td>
<td>29.2 4.1 (89)</td>
<td>29.9 3.0 (15)</td>
<td>25.9 4.4</td>
</tr>
<tr>
<td>Bennett Mechanical</td>
<td>37.8 10.3 (90)</td>
<td>32.5 9.1 (15)</td>
<td>33.1 9.2</td>
</tr>
<tr>
<td>Comprehension (AA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAT Clerical Speed</td>
<td>62.5 15.1 (66)</td>
<td>69.2 13.1 (14)</td>
<td>47.2 11.0</td>
</tr>
<tr>
<td>and Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concept Mastery Test</td>
<td>52.1 16.8 (67)</td>
<td>59.6 26.7 (14)</td>
<td>85.6 11.1</td>
</tr>
<tr>
<td>Coding</td>
<td>73.6 11.1 (32)</td>
<td>85.4 7.1 (8)</td>
<td>64.8 9.0</td>
</tr>
<tr>
<td>TSWE</td>
<td>50.2 8.8 (91)</td>
<td>53.5 6.7 (15)</td>
<td>55.6 4.5</td>
</tr>
<tr>
<td>Factor 1 (spatial/speed)</td>
<td>.27 (.72)</td>
<td>-.11 (.70)</td>
<td>.92 (.77)</td>
</tr>
<tr>
<td>Factor 2 (verbal)</td>
<td>-.28 (.81)</td>
<td>.62 (.41)</td>
<td>.99 (.47)</td>
</tr>
<tr>
<td>Factor 3 (nonverbal reasoning)</td>
<td>.107 (.80)</td>
<td>-.67 (.79)</td>
<td>.23 (.87)</td>
</tr>
</tbody>
</table>

Downloaded from gq.sagepub.com at PENNSYLVANIA STATE UNIV on May 8, 2016
on the Raven's. The Bennett Mechanical Comprehension Test proved to be slightly more difficult. Even so, the mean score of these students was comparable to the average earned by 12th-grade males. We conclude that these students' nonverbal aptitudes were highly developed.

Results on the verbal tests were similar. On the Concept Mastery Test the students, while still 13-years-old, scored slightly better than a special sample of Air Force captains (Terman, 1956). They also knew a great deal of English grammar, as demonstrated by their TSWE scores. On that test they scored at approximately the 80th percentile of college-bound 12th-graders (ATP, 1984). Finally, on the speed test the extremely precocious students scored at the 90th percentile of 12th-graders, and on the memory test their scaled (according to the WISC-R procedure) score was 16; 10 is considered average for their age.

**Talent Group Differences**

There were differences in pattern of performance between the two talent groups (see Table 1). The mathematically precocious students scored higher than the verbally precocious students on the spatial, nonverbal reasoning, speed, memory, and mechanical comprehension tests. In contrast, the 630V's scored higher than the 700M's on the verbal and general information test and the test of mechanics of English expression. All the differences between the groups were significant by t-tests, except for mechanical comprehension (p = .08). Moreover, all the associated effect sizes were large, except for mechanical comprehension (small). The highest mean scores, however, were primarily obtained by the Doubles: those 18 who met both the verbal and mathematics criteria.

As a check of the above conclusions, a regression analysis was performed for each test with group and sex as the independent variables. As expected, the important independent variable was talent group, which was significant beyond the .001 level for all tests except TSWE (p < .01) and the Bennett Mechanical Comprehension (n.s.). The independent variable sex and the sex by group interaction contributed little to the equations and were not significant.

**Relationship Among Test Performances**

A principal-axis factor analysis of the scores from eight of the tests was then performed. Excluded was Coding, which had been administered to too few students. Moreover, because the sample size was too small, analyses could not be performed separately by sex and group. Three factors emerged; they accounted for 68% of the variance. The factors were then rotated using the Oblimin method. The resulting pattern matrix is shown in Table 2. The first factor loaded on all the highly speeded tests: the three spatial ability measures and the clerical speed and accuracy test (our measure of speed). Cubes had the highest loading, followed by the accuracy test. Accordingly, the first factor was labeled spatial/speed. The second factor loaded on the TSWE and the Concept Mastery Test. We labeled it as a verbal factor. The third factor loaded on the various nonverbal reasoning tests, especially the Bennett Mechanical Comprehension and Raven's Progressive Matrices. We identified this factor as nonverbal reasoning.

**Table 2**

Rotated Factor Matrix of the Three Factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Visualization</td>
<td>.40</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>Cubes</td>
<td>.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Orientation</td>
<td>.40</td>
<td>.77</td>
<td>.53</td>
</tr>
<tr>
<td>Mechanical Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raven's P.M.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clerical Speed &amp; Accuracy</td>
<td>.56</td>
<td>.74</td>
<td>.77</td>
</tr>
<tr>
<td>Concept Mastery Test</td>
<td>.74</td>
<td>.77</td>
<td></td>
</tr>
</tbody>
</table>

Factors 1 and 3 correlated .41. Three of the four tests that loaded on factor 1 (i.e., the spatial/speed factor) were spatial ability measures, which is usually considered a nonverbal ability (our factor 3). This probably accounts for the relationship between factors 1 and 3. No other substantial correlations between factors were found (i.e., .02 and .08).

Factor scores were then computed for each individual (Table 1). ANOVAs by group (630V vs. 700M) and sex were then performed on the factor scores. The analyses showed that, for each factor, talent group was the significant variable (p < .001 for the three factors). Sex and the group by sex interaction were not important. The performance of extremely mathematically talented students was superior to that of the extremely verbally talented students on the spatial/speed and nonverbal reasoning factors. For the verbal factor the verbally talented students exhibited higher performance.

Finally, an interesting trend was revealed. The presence of exceptionally high verbal ability appeared to increase the likelihood of the presence of high mathematical ability. Only one of the verbally precocious students had an SAT-M score lower than 500 (the average score of a college-bound 12th-grade male). The reverse was not apparent: high mathematical ability did not seem to indicate concomitantly high verbal ability. Twenty-two students scoring 700 or above on the SAT-M scored below 430 on the SAT-V (the average score of a college-bound 12th-grade male). These results, which were significantly different (p < .05), indicate that verbally precocious students may be more evenly balanced in their cognitive profiles than mathematically precocious students.

**Discussion**

The construct, giftedness, is poorly defined and understood (Hoge, 1988), a fact which has contributed to existing
problems in the way students are identified as gifted (Feldhusen, 1989). It is now believed that giftedness should be viewed as comprised of multiple talents rather than one general ability (e.g., Gagne, 1985; Gardner, 1983; Stanley, 1984b). Is such a view justifiable? We investigated differences in the pattern of special cognitive abilities of extremely verbally precocious students compared to extremely mathematically precocious students. The two forms of giftedness were found to be distinct, a finding which is consistent with viewing giftedness as comprised of several different talents.

Verbally and mathematically precocious students differed in their patterns of cognitive strength on the individual tests and the factors of which those tests are comprised. Not surprisingly, the verbally precocious scored higher on verbal and general knowledge types of tests, and the strengths of the mathematically precocious were in the nonverbal abilities. Moreover, we factor analyzed the scores on the various ability tests to identify the model of intelligence that best fit the data (e.g., a single factor, "g," or multiple talents). Three factors were identified by the factor analysis. They were labeled spatial/speed, verbal, and nonverbal reasoning. These results are somewhat compatible with Horn and Cattell's (1966) crystallized and fluid intelligence model. Crystallized intelligence is heavily dependent on culturally loaded, fact-oriented learning. Tasks highly correlated with this factor include vocabulary and general information. In contrast, fluid intelligence demands little in the way of specific informational knowledge. It reflects the ability to see complex relationships. Our data and those of Benbow et al. (1983) and Pollins (1984) were consistent with such a dichotomization of intelligence or giftedness.

Our results also indicate that speed might be an important component of extreme giftedness. Dark and Benbow (in press; submitted) present evidence indicating that enhanced working memory and speed were components of giftedness, but were more clearly aligned with mathematical than with verbal talent. The data presented in this paper, which were obtained using a psychometric research paradigm, converge with those obtained by Dark and Benbow, who used an information processing approach to investigate the giftedness construct. Enhanced memory and speed appear to be associated more strongly with mathematical than with verbal talent. Further studies are needed to confirm the importance of the speed factor, however. In our study the spatial/speed factor may be artificial because the spatial tests in our test battery were highly speeded and the spatial/speed and nonverbal reasoning factor correlated highly.

Becker (1989) had found that among mathematically able students spatial ability did not relate to item performance on the SAT-M. In this study, however, we found that mathematical talent was associated with spatial and nonverbal abilities, as did Cohn (1977). Burnett, Lane, and Dratt (1979) and others have presented similar relationships, but among average-ability students.

Gagne (1985) suggested that abilities interact with environmental conditions and personality factors to emerge as talent in a specific domain. In a series of studies, we now have compared these verbally and mathematically precocious students along several dimensions. In support of Gagne's model, we previously found some differences in personality traits (Brody & Benbow, 1986; Dauber & Benbow, in press) and in environmental circumstances (i.e., an emphasis on books and reading; Benbow, 1989). Now we also have identified differences in specific cognitive abilities. Mathematical talent and verbal talent do appear, therefore, to relate to a different mix of cognitive abilities, personality traits, and environmental circumstances.

No gender differences on any of the specific aptitude tests were found, even though several of the measures utilized traditionally exhibit gender differences. Although based on a small sample size of females, these results suggest that when boys and girls are selected by the same ability criteria, their profiles of specific aptitudes are comparable. Thus, contrary to our predictions, it is unlikely that differences in underlying abilities between boys and girls can explain gender differences in mathematics and science achievement among mathematically precocious students.

This study is limited by the small sample size and the extremely rare sample of gifted students tested. The structure of intelligence among gifted students with less exceptional abilities may differ.

Implications for Identification

Our results indicate that giftedness is not a unitary construct. Verbal and mathematical precocity are distinct forms of intellectual giftedness; they are associated with different cognitive profiles. Thus, students should be selected for special academic programs based upon qualities required by that program. Selecting students for a "gifted" program on the basis of one overall ability does not appear justifiable. Evidence presented in this paper suggests that such a procedure might exclude too many nonverbally gifted students from "gifted" programs. Such students are less balanced in their specific abilities and, therefore, may not rank high on an overall or combined index of ability. This point needs further study.

The results in this paper also provide some support for the use of the Raven's Progressive Matrices Test for identification of mathematically gifted students. Raven's test performance was high for extremely gifted students but was more closely aligned to mathematical than verbal precocity. Moreover, Benbow (in preparation) found that Raven's test scores predicted performance of gifted students in fast-paced mathematics classes.

In conclusion, this study provided a unique opportunity to study the cognitive abilities of students selected as being extremely verbally and/or mathematically precocious. The pattern of performance on tests of the specific cognitive abilities
commonly associated with intelligence differed for verbal compared to mathematical precocity. This difference provides justification for a multiple-talents approach to identification of the gifted. Gifted students should be selected for special programs on the basis of having qualities that match the intent of the program, not on the basis of one overall ability. In addition, programs for the gifted should be designed to serve the associated educational needs of the two types of giftedness identified in this study.

Author's Note: We thank Linda J. Allred, Linda E. Brody, Keith Davis, Bert F. Green, Douglas C. Langfeld, Julian C. Stanley, and Mary Ann C. Swiatek for helpful comments and suggestions. Financial support was provided by a NSF grant to Camilla P. Benbow (MDR-8651737).

References