DOCUMENT RESUME

ED 391 674 SE 057 681

AUTHOR	Stanley, Julian C.; And Others
TITLE	Ipsative Evaluative Attitudes versus Vocational Interests and Cognitive Abilities of Bright Male Versus Female Seventh-Graders.
PUB DATE	Apr 95
NOTE	41p.; Paper presented at the Annual Meeting of the American Educational Research Association (April 20, 1995).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academically Gifted; *Cognitive Ability; Grade 7; Junior High Schools; *Junior High School Students; *Sex Differences; *Student Interests; Tests; *Values

ABSTRACT

This empirical study involves three types of "tests": values, interests, and cognitive abilities useful for learning science. It also involves three methods of test-battery construction or use: ipsative (forced-choice), non-ipsative ("normative"), and ipsatived (i.e., normative scores changed to ipsative ones). Scores from ipsatively constructed or ipsatively scored test batteries have no overall level; every examinee earns the same total score. Thus, ipsativity creates strange statistics, but sometimes (as in this study) yields interesting results. For the bright seventh-grade students studied, some of the main findings are as follows: (a) The boys' Theoretical evaluative attitude relates best to the eight cognitive tests, whereas the girls' Aesthetic evaluative attitude relates best; (b) The boys are much more Theoretical, Economic, and Political than the girls, and the girls are much more Aesthetic, Social, and Religious than the boys (the same findings as for college students); (c) The cognitive test scores intercorrelate surprisingly highly, considering the great selectivity of the samples studied; and (d) There is considerable agreement between scores on the intrinsically ipsative Allport-Vernon-Lindzey Study of Values and those on the normative Holland Occupations Checklist (HOC), and even more so when the HOC scores are forced to become ipsative. (Author)

Ipsative Evaluative Attitudes Versus Vocational Interests and Cognitive Abilities of Bright Male Versus Female Seventh-Graders

Julian C. Stanley and Heinrich Stumpf
Johns Hopkins University

Sanford J. Cohn
Arizona State University

Abstract

This empirical study involves three types of "tests": values, interests, and cognitive abilities useful for learning science. It also involves three methods of test-battery construction or use: ipsative (forced-choice), non-ipsative ("normative"), and ipsatived (i.e., normative scores changed to ipsative ones). Scores from ipsatively constructed or ipsatively scored test batteries have no overall level; every examinee earns the same total score. Thus, ipsativity creates strange statistics, but sometimes (as in this study) yields interesting results. For the bright seventh-grade students studied, some of the main findings are as follows: the boys' Theoretical evaluative attitude relates best to the eight cognitive tests, whereas the girls' Aesthetic evaluative attitude does; the boys are much more Theoretical, Economic, and Political than the girls, and the girls are much more Aesthetic, Social, and Religious than the boys (the same findings as for college students): the cognitive test scores intercorrelate surprisingly highly, considering the great selectivity of the samples studied; and there is considerable agreement of scores on the intrinsically ipsative Allport-Vernon-Lindzey Study of Values with those on the normative Holland Occupations Checklist, and even more so when the HOC scores are forced to become ipsative.

Ipsative Evaluative Attitudes Versus Vocational Interests and Cognitive Abilities of Bright Male Versus Female Seventh-Graders

Julian C. Stanley* and Heinrich Stumpf — Johns Hopkins University
Sanford J. Cohn — Arizona State University

This may seem to be a strange study. Why should anyone analyze data from an ipsative self-report value "test" designed for college students and adults, last standardized before 1960 and administered in 1977 to bright seventh-graders? In the first place, self-report devices usually don't match ability tests with respect to reliability and validity.

Secondly, ipsatively constructed test batteries have peculiar psychometric properties (Gleser, 1972). They yield several subtest scores, but every examinee has exactly the same total score (Cattell, 1944; Gordon, 1951, 1972; Clemans, 1966; Hicks, 1970; McDermott et al., 1992). Then *n* subtest scores inter-correlate mostly negatively, the mean *r* being $-1/(n - 1)$ or a *little* larger.

Thus, only the shape and dispersion of an examinee's subtest scores are supposed to be meaningful. Strictly speaking, creating a distribution of a given subtest's scores across individuals is illegitimate. Yet, curiously, inter-individual comparisons are usually made and often yield means, standard deviations, correlation coefficients, and gender differences that make sense (e.g., Linsenmeier, 1976). Unexpected results can occur, as in this report, and yet appear to be somewhat interpretable.

Thirdly, it may seem foolish to try to assess the values of bright 12- and 13-year-old boys and girls with an adult-level scale. Do they know enough to respond appropriately to its ideas? Have their evaluative attitudes matured sufficiently?

Despite these seemingly severe restrictions, we have taken data from Sanford J. Cohn's Master's-degree paper (Cohn, 1978) and analyzed it further. Then we have been able to check part of it against much more recent data (Lubinski and Benbow, 1992, 1994; Lubinski, Benbow, and Sanders, 1993; Mills and Stumpf, personal communication).[1]

Our aim is exploratory, to generate hypotheses rather than test them. What we found confirms for mathematically able 12- and 13-year-old students the pattern of gender differences noted since 1930 for college students and adults on the Allport-Vernon (later, in 1951 and 1960, the Allport-Vernon-Lindzey) Study of Values (Vernon and Allport, 1931; Stanley and Waldrop, 1952; Allport, Vernon, and Lindzey, 1970). That is interesting, but not the most novel of the results we are presenting here.

Objectives

We explore use of the intrinsically ipsative Allport-Vernon Lindzey "Study of Values" (SV) and the normative (i.e., non-ipsative) "Holland Occupations Checklist" (HOC) with bright young seventh-graders, mostly aged 12 or 13. These instruments, quite different in content and underlying theory, nevertheless tap some of the same constructs when applied to adults. They have been found useful in vocational and educational counseling. By relating them to each other and to eight diverse ability tests, we further assess their utility for studying evaluative attitudes and vocational interests of gifted children.

Theoretical Framework

The SV was originally based on social philosopher Eduard Spranger's Lebensformen (1927) (translated as Types of Men, 1928), which posits that each person's life is a succession of forced choices guided by six competing evaluative attitudes: Aesthetic (interest in form and harmony), Economic (usefulness), Political (power), Religious (mystical unity), Social (love of people), and Theoretical (discovery of truth). Famed Harvard social psychologist Gordon Allport and British psychometrician Philip Vernon used Spranger's theory to construct a values inventory that pits each of the evaluative attitudes systematically against every other, thereby constructing a set of six interdependent scales (Vernon and Allport, 1931). A revision containing 15 items per scale appeared in 1951, and another in 1960 (see Figure 1). High school norms were published in 1970.

(Please put Fig. 1 about here.)

The overall score of each examinee is the same as that of any other examinee, so every score profile is "ipsative" (from the Latin <u>ipse</u>, self). One must "borrow from Peter to pay Paul." For example, "Whom do you admire more, Marie Curie or Florence Nightingale?" Choose Curie and earn a point on the Theoretical scale, whereas choosing Nightingale gives a point on the Social (service) scale.

By contrast, psychologist John L. Holland's "Self-Directed Search," from which the HOC is taken, is based on his carefully-worked-out theory of occupations (Holland, 1985). The checklist (Holland, 1970) consists of six categories (Artistic, Conventional, Enterprising, Investigative, Realistic, and Social), each containing 14 occupations. There are no restraints on how many or how few occupations within a category the examinee may choose, from 0 to 14. Thus, an individual's total score over the 6 categories may be as low as 0 or as high as 84. Each profile has shape and dispersion (which the SV also does) and level (which the SV does not).

Although the HOC is non-ipsative, its raw scores can be made "ipsative" in a variety of ways. Such ways tend to yield various results. We chose the simplest, within-examinee, method by dividing an examinee's score for each category by his or her total score. Then every examinee's overall score becomes the same, namely 1.00. Intuitively, ipsativing the HOC in this way might seem likely to make it perform more like the SV. Most research has found, however, that ipsativing across examinees does not yield much new information (e.g., McDermott et al. 1992), but perhaps never before have results of ipsativing within examinees been compared, across a variety of statistics, with those of an intrinsically ipsative instrument.

Tests and Subjects

The SV, HOC, and eight non-ipsative aptitude and achievement tests (Scholastic Aptitude Test--Mathematical, SAT-M; SAT-Verbal; American College Testing Program--Mathematics, ACT-M; ACT-Natural Science, ACT-NS; Differential Aptitude Test--Abstract Reasoning, DAT-AR; DAT-Mechanical Reasoning, DAT-MR; DAT-Spatial Relations, DAT-SR; and Cooperative Mathematics Test Algebra I, Alg. I, alternate Form B) were administered to 188 male and 90 female seventh-graders. They had been the top-scoring third in a three-state mathematics talent search in the East in 1977. The selection criterion was $2(\text{SAT-M}) + 1(\text{SAT-V}) \geq 1330$. Boys qualified more easily via M, girls via V.

Statistical Analysis

For each examinee there were the following 26 scores: 6 on SV, 6 on non-ipsative HOC, 6 on ipsatived HOC, and 8 on the ability tests. Intercorrelations of these scores were computed, separately by sex. Also, a gender-difference effect size[2] was computed for each of the 26 measures. In 4 other (six-by-six-matrix) comparisons, special attention was paid to the correspondence of the SV with the non-ipsative HOC versus the ipsatived HOC, separately by sex.

Correlation coefficients could not be expected to be high because of a number of reasons: the great degree of selection of the subjects, probably causing restriction of range of scores (but difficult tests helped spread out the scores); the ipsative nature of the Study of Values and of the ipsatived Holland Occupations Checklist; and the brevity of the SV scales (15 items each) and the HOC scales (14 items each). We had no direct estimates of reliability or internal-consistency coefficients, but comparison of the standard deviations of the SV scores with those in the SV Manual (Allport, Vernon, and Lindzey, 1970) revealed fairly close agreement, even though those norms were obtained from extremely heterogeneous groups of college students across the entire United States. Ratios of the variance of same-sex norm group scores to the variance of scores of our seventh-graders ranged from 1.73 for boys on Economic (i.e., standard deviations of 7.92 vs. 6.03) to 0.74 for girls on Theoretical (7.19 vs. 8.33). Because our subjects were exceptionally bright, conscientious test takers, it seems likely that their standard errors of measurement (Stanley, 1971) were less than that of the norm group, and therefore that the internal consistency of their SV scores would not be greatly smaller than the substantial coefficients (.84-.95, mean .90) reported in the manual. This inference is reinforced by the sizable intercorrelations for these subjects on the eight ability tests over a three-month interval, as high as .74 between two different mathematics tests, despite great explicit restriction of score range on one of them.

Results

(1) Effect sizes for SV (see Table 1) favoring boys were 1.01 standard

(Put Table 1 about here.)

deviations for Theoretical, 0.95 for Economic, and 0.56 for Political; favoring girls were -1.13 for Aesthetic, - 0.69 for Social, and -0.48 for Religious. These agree substantially with results for college students we computed from the SV Manual (Allport, Vernon, and Lindzey, 1970): 1.10, 0.65, 0.79, -0.90, -0.70, and -0.60, respectively. They are also similar to SV effect sizes for Camilla Benbow and David Lubinski's recent study of bright young boys and girls attending their academic summer programs in Iowa: 0.81, 0.69, 0.62, -0.75, -0.73, and -0.37 (Lubinski and Benbow, 1992).

The effect sizes found by Carol P. Mills and Heinrich Stumpf (personal communication, 21 March 1995) for bright participants in academic summer programs in the East, similar to Benbow's in Iowa, were as follows: T 0.41. E 0.79, P 0.92, A -0.87, S -0.51, and R -0.40. Thus, all three sets of data, from different times, places, and selection procedures, agree fairly well. Clearly, on the SV there is the usually found substantial gender difference: Theoretical, Economic, and Political favor males, whereas Aesthetic, Social, and Religious favor females: (e.g., see Lubinski, Schmidt, and Benbow, Under review).

For the HOC, non-ipsative and (ipsatived), the effect sizes tended to be smaller (see Table 2): favoring males, 0.41 (0.67) for Realistic, 0.15

(Put Table 2 about here.)

(0.52) for Conventional, 0.10 (0.18) for Investigative, and 0.00 (0.35) for Enterprising; favoring females, -0.91 (-0.88) for Social and -0.79 (-0.65) for Artistic. The largest discrepancy was for HOC Investigative versus SV Theoretical, perhaps due to the two types of occupations in the HOC I category. Some are "hard science" or engineering, others "soft science," writing, or editing. Males probably tend to choose the former, females the latter.

The corresponding Mills and Stumpf HOC (non-ipsative) effect sizes were R 0.24, C 0.26, I -0.30, E 0.02, S -0.20, and A -0.53.

Seven of the eight effect sizes for the ability tests (see Table 3)

(Put Table 3 about here.)

favored the boys: 0.73 (SAT-M), 0.72 (DAT-MR), 0.62 (ACT-M), 0.29 (ACT-NS), 0.26 (Alg. I), 0.11 (DAT-AR) and 0.07 (DAT-SR). SAT-Verbal favored the girls (-0.16). In Benbow and Lubinski's groups, the largest ability-test effect size was 0.97 for mechanical comprehension, favoring boys.

(2) The four correlation-matrix comparisons of HOC with SV (Tables 4-7) show that, although measuring somewhat different constructs, they have

(Put Tables 4, 5, 6, and 7 about here.)

much in common: Artistic correlates better with Aesthetic (.42, .41, .39, .33) than with any of the 10 other possibilities, Investigative with Theoretical, Social with Social, and Conventional with Economic. Patterns for the other two interests and two values were sex-related and less clear. As we expected, the ipsatived HOC related somewhat more strongly to SV, which is intrinsically ipsative, than did the non-ipsative HOC.

3) All eight ability tests (see Table 8) intercorrelated positively

(Put Table 8 about here.)

for the boys, from .74 for ACT-M with SAT-M to .01 for Algebra I with DAT-MR. The 28 *r*s for the girls, four of which were negative, ranged from .56

for ACT-M with SAT-M to -.14 for SAT-V with DAT-MR.

4) A striking result was found. For the boys (see Table 9), SV

(Put Table 9 about here.)

Theoretical correlated positively with each of the ability tests, whereas Political correlated negatively with each. For the girls (see Table 10),

(Put Table 10) about here.)

the "cognitive" SV was Aesthetic, and there was no consistently negative relationship. This gender difference seems fairly large. It is based on correlation coefficients--rather small ones--so only associative, not causal, inferences are warranted. As noted, Theoretical scores of the boys corresponded decidedly best with their ability scores, whereas Aesthetic scores corresponded decidely best for the girls. Why? Will this difference be found among adults? Does it have implications for the teaching of, for example, physics? That is, should the instruction for boys be primarily theoretical and analytical, but the instruction of girls emphasize aesthetic aspects of the subject? Some feminists would, on intuitive grounds, say so.

The ongoing work of Eccles, Wigfield, Harold, and Blumenfeld (1993) is also relevant. They found that "Boys had . . . more positive competence beliefs for mathematics. Girls had more positive competence beliefs and values than boys did for reading and music activities." From other sources there is much evidence that boys tend to <u>reason</u> better mathematically than girls, and of course that girls tend to be considerably more aesthetic than boys and also somewhat better at language usage (e.g., Benbow, 1988, 1990; Stanley et al., 1992; Stanley, 1993; Stumpf and Stanley, under review).

Mills (1981, 1993), found a "different relationship for males and females . . . between high intellectual scores and indicators of self-concept and emotional well-being" and intriguing relationships between scores on the Myers-Briggs Type Indicator and ability-test scores of mathematically talented boys versus girls.

(5) For the boys (see Table 11), HOC Investigative had the strongest

(Put Table 11 about here.)

positive relationship to ability-test scores, whereas Social related most negatively to them. The ability test most closely related to the HOC categories was DAT-MR. For girls (see Table 12), the largest positive rs

(Put Table 12 about here.)

were for Conventional, the largest negative for Social and Realistic. The ability most closely related to their HOC interests was ACT-M. Ipsativing did not result in different categorization, but strengthened the relationship (see Tables 13 and 14).

(Put Tables 13 and 14 about here.)

(6) SV intercorrelations are interesting. For the extensive Iowa data the 15 rs average -.191 for boys and -.194 for girls. The largest rs are between Economic and Political, .26 for boys and .25 for girls. These seem small, but as deviations from the respective means they become .45 and .44 (see Table 15). The most-negative rs are between Religious and

(Put Table 15 about here.)

Theoretical, -.49 for the boys and -.46 for the girls. As deviations from the respective means these are -.30 and -.26.

Aesthetic correlates with Theoretical -.10 for the boys (deviation, .09) and -.25 for the girls (deviation, -.06). This suggests faintly that more boys than girls may score highest on both Aesthetic and Theoretical, the presumably creative dyad. That is inconsistent, however, with data for college students in the SV Manual, where r = -.10 for males and .07 for females.

All 30 intercorrelations of the non-ipsative HOC categories are positive, ranging from r = .16 to .63. The median r for the males is .38, for the females .49. Ipsatived HOC categories exhibit the same pattern of mainly negative intercorrelations as SV categories do.

Discussion

From this abundance of data it becomes obvious that SV and HOC, especially when ipsatived, produce interesting relationships even though administered to bright boys and girls far younger than the populations for which they were designed. SV and HOC can be used to analyze the development of interests and attitudes, as Benbow and Lubinski are doing. Especially importantly, their predictive power for foretelling later achievements is being studied longitudinally by them. MacKinnon (1962) and Helson (1980) found both high theoretical and high aesthetic values among creative adults. Also, we of SMPY have noted informally that youth of either sex who go on to achieve well in research tend to have scored, at age 12, high on Theoretical.

How do values and interests develop from age 12 or 13 until adulthood? Which grow the most? Lubinski, Benbow, and Ryan (In press) found appreciable stability of vocational interests of bright boys and girls from adolescence to adulthood. Lubinski, Schmidt, and Benbow (Under review) studied change in SV scores from age 12 or 13 to 32 or 33. Marsh (1989) found that "Sex differences in specific areas of self-concept were generally consistent with sex stereotypes, and relatively stable from preadolescence to early adulthood." McCrae and Costa (1994) provide evidence for the stability of adult personality. Over a two-year period, Terwilliger and Titus (1995) found gender differences and, also, decline for both sexes in most of their scales assessing "components of attitude thought to be relevant to success in a program for mathematically talented youth . . ."

Does ipsativity produce less-stable individual profiles than non-ipsativity? Logically, perhaps it should. If an ipsatively measured value or interest increases with age, which other values decline? Does a non-ipsative battery such as the HOC predict adult achievement less well than when these scores are ipsatived? Why are some interests and evaluative attitudes more cognitively loaded than others? Has this confused researchers, because seldom are both self-report scores and ability-test scores considered together?

This study has explored relationships among interests, evaluative attitudes, and abilities in bright seventh-grade boys versus girls. Its findings might outline a fertile, neglected field for further research, especially of a predictive nature. It may have implications for selecting Westinghouse Science Talent Search scholarship awardees and persons to attend the no-cost annual Research Science Institute and other programs, such as residential state high schools, oriented toward encouraging able young people to become scientists. Ability measures alone, including grades in school, may screen into such programs a number of bright students who have little potential for research, and screen out others who do. Reducing the number of such false positives and false negatives would enhance the effectiveness of those programs. It is difficult, however, to use self-report devices such as SV and HOC for selection, because scores on them can readily be faked.

A number of the gender differences found for these youth are large. Why do boys and girls have such different evaluative attitudes, interests, and abilities? Do few of either sex have "androgynous" profiles (Lubinski, Tellegen, and Butcher, 1981), high on both Theoretical and Aesthetic evaluative attitudes? Does that mean, as MacKinnon's and Helson's work seems to suggest, that few of them will become highly creative architects or scientists? Or do profiles tend to move with age toward more high-T high-A scorers? Systematic longitudinal studies of the kind that Lubinski, Schmidt, and Benbow (Under review) are conducting can answer some of these questions.

References

Allport, G. W., Vernon, P. E., & Lindzey. G. (1970). Manual for the Study of Values (3rd ed.). Boston: Houghton Mifflin.

Benbow, C. P., (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. Behavioral and Brain Sciences, 11, 169-232.

Benbow, C. P. (1990). Sex differences in mathematical reasoning ability: Further thoughts. Behavioral and Brain Sciences, 13, 196-198.

Cattell, R. B. (1944). Psychological measurement: Normative, ipsative, interactive. Psychological Review, 51, 292-303.

Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. Psychometric Monographs, 14.

Cohn, S. J. (1978). Toward a national talent search. Master's paper, Department of Psychology, Johns Hopkins University, Baltimore, MD.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and and gender differences in children's self- and task perceptions during elementary school. Child Development, 64, 830-847.

Gleser, L. J. (1972). On the bounds of the average correlation between subtest scores in ipsatively scored tests. Educational and Psychological Measurement, 32, 759-766.

Goodwin, W. L. (1964). Adjustment for sex and variability differences on the Allport-Vernon-Lindzey Study of Values profiles. Journal of Educational Measurement, 1, 55-58.

Gordon, L. V. (1951). Validities of the forced-choice and questionnaire methods of personality measurement. Journal of Applied Psychology, 35, 407-412.

Gordon, L. V. (1972). A typological assessment of "A Study of Values" by Q-Methodology. Journal of Social Psychology, 86, 55-67.

Helson, R. (1980). The creative woman mathematician. In L. H. Fox, L. Brody, & D. Tobin (Eds.), Women and the mathematical mystique (pp. 23-54). Baltimore, MD: Johns Hopkins University Press.

Hicks, L. E. (1970). Some properties of ipsative, normative, and forced normative measures. Psychological Bulletin, 74, 167-184.

Holland, J. L. (1970). Holland Occupations Checklist. From The Self-Directed Search. Tampa, FL: Consulting Psychologists Press.

Holland, J. L. (1985). Making vocational choices: A theory of vocational personalities and work environments (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Linsenmeier, J. A. W. (1976). Random vs. nonrandom Study of Values profiles. In D. P. Keating (Ed.), Intellectual talent: Research and development (pp. 285-292). Baltimore, MD: Johns Hopkins University Press.

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math-science pipeline. Current Directions in Psychological Science, 1, 61-66.

Lubinski, D., & Benbow, C. P. (1994). The Study of Mathematically Precocious Youth: The first three decades of a planned 50-year study of intellectual talent. In R. F. Subotnik & K. D. Arnold (Eds.), Beyond Terman: Contemporary longitudinal studies of giftedness and talent (pp. 255-281). Norwood, NJ: Ablex Publishing Co.

Lubinski, D., Benbow, C. P., & Ryan, J. (In press). Stability of vocational interests among the intellectually gifted from adolescence to adulthood: A 15-year longitudinal study of Holland's vocational-interest themes ("RIASEC"). Journal of Applied Psychology.

Lubinski, D., Benbow, C. P., & Sanders, C. E. (1993). Reconceptualizing gender differences in achievement among the gifted. In K. A. Heller, F. J. Monks, & A. H. Passow (Eds.), International handbook of research and development of giftedness and talent (pp. 693-707). Oxford, England: Pergamon Press.

Lubinski, D., Schmidt, D. B., & Benbow, C. P. (Under review). Stability of evaluative ("personological") attitudes among the intellectually gifted from adolescence to adulthood: A 20-year longitudinal study. Ames, IA 50011: The authors, Department of Psychology, Iowa State University.

Lubinski, D., Tellegen, A., & Butcher, J. N. (1981). The relationship between androgyny and subjective indicators of emotional well-being. Journal of Personality and Social Psychology, 40, 722-730.

MacKinnon, D. W. (1962). The nature and nurture of creative talent. American Psychologist, 17, 484-495.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. Journal of Educational Psychology, 81, 417-430.

McCrae, R. R., & Costa, Jr., P. T. (1994). The stability of personalities: Observations and evaluations. Current Directions in Psychological Science, 3, 173-175.

McDermott, P. A.; Fantuzzo, J. W.; Glutting, J. J.; Watkins, M. W.; & Baggaley, A. R. (1992). Illusions of meaning in the ipsative assessment of children's ability. Journal of Special Education, 25, 504-526.

Mills, C. J. (1981). Sex roles, personality, and intellectual abilities in adolescents. Journal of Youth and Adolescence, 10, 85-112.

Mills, C. J. (1993). Personality, learning style and cognitive style profiles of mathematically talented students. European Journal for High ability, 4, 70-85.

Spranger, E. (1927). Lebensformen. Halle (Saale): Max Niemeyer.

Spranger, E. (1928). Types of men: The psychology and ethics of personality. (Translation of Lebensformen). Halle (Saale): Max Niemeyer.

Stanley, J. C. (1971). Reliability. In R. L. Thorndike (Ed.), Educational Measurement (2nd ed.) (pp. 356-442). Washington, D.C.: American Council on Education.

Stanley, J. C. (1993). Boys and girls who reason well mathematically. In G. R. Bock & K. Ackrill (Eds.), The origins and development of high ability (pp. 119-138). New York: Wiley.

Stanley, J. C.; Benbow, C. P.; Brody, L. E.; Dauber, S. L.; & Lupkowski, A. E. (1992). Gender differences on eighty-six nationally standardized aptitude and achievement tests. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), Talent Development, Vol. 1. Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development (pp. 42-65). Unionville, NY: Trillium Press.

Stanley, J. C., & Waldrop, R. S. (1952). Intercorrelations of Study of Values and Kuder Preference Record scores. Educational and Psychological Measurement, 12, 707-719.

Stumpf, H., & Stanley, J. C. (Under review). Gender-related differences on the College Board's Advanced Placement and Achievement Tests, 1982-1992.

Terwilliger, J. S., & Titus, J. C. (1995). Gender differences in attitudes and attitude changes among mathematically talented youth. Gifted Child Quarterly, 38, 29-35.

Vernon, P. E.; & Allport, G. W. (1931). A test for personal values. Journal of Abnormal and Social Psychology, 26, 231-248.

Footnotes

*Presented by Dr. Stanley on 20 April 1995 in San Francisco at the annual meeting of the American Educational Research Association. Address inquiries to Professor Julian C. Stanley, SMPY, 156A Bloomberg Center, Johns Hopkins University, Baltimore, MD 21218, telephone (410)516-6179, fax (410)516-7239, and e-mail setcty@jhunix.hcf.jhu.edu . We thank Camilla Benbow, William V. Clemans, Leonard V. Gordon, Samuel A. Livingston, David Lubinski, and Carol J. Mills for helpful suggestions.

[1] We thank David Schmidt and Babette Suchy of Iowa State University for computing for us a number of statistics. Over the years from 1976 onward, many others have also contributed to the various aspects of this study.

[2] $(\text{Mean}_{\text{boys}} - \text{Mean}_{\text{girls}}) / \sqrt{\dfrac{\text{Variance}_{\text{boys}} + \text{Variance}_{\text{girls}}}{2}}$.

Figure 1

Allport-Vernon-Lindzey (1960) Norms, National Standardization Population (2489 Men, 1289 Women), Plotted According to Normalized Standard Scores. Profile of a High-Theoretical, Low-Economic, Low-Political Man. Also Note the Percentile Ranks for a Woman Earning Those Scores. See Goodwin (1964).

	Centile	Theoretical M	Theoretical F	Economic M	Economic F	Aesthetic M	Aesthetic F	Social M	Social F	Political M	Political F	Religious M	Religious F	Centile
HIGH	99	60-62	52-54	60-63	54-56	54-56	61-64	53-55	58-60	58-60	52-53	59-62	65-68	99
	98	59	50-51	59	53	52-53	60	51-52	56-57	56-57	51	57-58	63-64	98
	97	58	49	58	52	51	58-59	50	55	55	50	56	61-62	97
	96	57	48	57	51	50	57		54		49	55	60	96
	94-95	55-56	47	55-56	49-50	48-49	56	48-49	53	53-54	48	53-54	59	94-95
	93						55				47	52	58	93
	92	54	46	54	48	47		47	52			51	57	92
	90-91	53	45	53		46	54	46	51	52	46	50	56	90-91
	88-89			52	47	45	53			51		49	55	88-89
	86-87	52	44		46		52	45	50	50	45		54	86-87
	83-85	51	43	51	45	44	51	44	49		44	47-48	53	83-85
AVERAGE	81-82		42	50		43	50			49			52	81-82
	78-80	50		49	44	42	49	43	48	48	43	46	51	78-80
	75-77	49	41	48	43	41		42	47		42	45		75-77
	71-74	48	40		42	40	48	41	46	47		44	49-50	71-74
	68-70			47		39	47			46	41	43		68-70
	64-67	47	39	46	41		46	40	45		40	42	48	64-67
	60-63	46	38	45	40	38	45	39	44	45		41	47	60-63
	56-59	45	37	44	39	37	44			44	39	40	46	56-59
	52-55					36	43	38	43			39	45	52-55
	49-51	44	36	43	38	35		37	42	43	38	38	44	49-51
	45-48	43	35	42	37	34	42				37	37	43	45-48
	41-44	42		41			41	36	41	42		36	42	41-44
	37-40		34		36	33	40	35	40	41	36	35	41	37-40
	33-36	41	33	40	35	32	39	34	39	40		34	40	33-36
	30-32	40	32	39	34	31					35		39	30-32
	26-29	39		38		30	38	33	38	39	34	33	38	26-29
	23-25		31	37	33	29	37	32	37	38		32	37	23-25
	20-22	38	30	36	32	28	36		36		33	31	36	20-22
	18-19	37					35	31		37		30		18-19
LOW	15-17	36	29	35	31	27	34	30	35	36	32	29	34-35	15-17
	13-14		28	34	30	26			34		31	28	33	13-14
	11-12	35	27	33	29	25	33	29		35		27		11-12
	9-10	34	26	32	28	24	32	28	33	34	30	26	31-32	9-10
	8					23	31	27	32		29	25		8
	7	33	25	31	27					33			30	7
	5-6	32	24	30	26	21-22	29-30	26	31	32	28	23-24	28-29	5-6
	4	31	23	29	25	20	28	25	30	31	27	22	27	4
	3	30	22	28	24	19	27	24	29	30	26	20-21	26	3
	2	28-29	21	26-27	23	17-18	25-26	22-23	27-28	29	25	18-19	24-25	2
	1	25-27	18-20	23-25	20-22	14-16	22-24	19-21	24-26	26-28	22-24	15-17	20-23	1

Table 1

Gender differences on the six scales of the Allport-Vernon-Lindzey

Study of Values among 188 boys and 90 girls

		Theoretical	Economic	Political	Religious	Social	Aesthetic
Males	Mean	48.21	42.82	44.72	31.31	38.36	34.62
	S.D.	6.74	6.03	6.61	9.58	5.88	6.99
Females	Mean	40.82	36.91	41.06	35.91	42.54	42.87
	S.D.	8.33	6.56	6.46	9.60	6.39	7.88
t (276d.f)		7.91	7.43	4.36	-3.74	-5.40	-8.83
Effect size (*d*)		1.01*	0.95	0.56	-0.48	-0.69	-1.13
p		$<.001$	$<.001$	$<.001$	$<.001$	$<.001$	$<.001$
d for 12-13-year-olds in recent Iowa State University Study, 395 boys and 327 girls		0.81	0.69	0.62	-0.37	-0.73	-0.75

*Was 1.10 in the latter 1950s for the Study of Values norm group of college students.

Table 2

Gender differences on the six scales of the *Holland Occupations Checklist* among 188 boys and 90 girls

		Realistic	Investigative	Conventional	Enterprising	Artistic	Social
Males	Mean	3.26	6.28	3.11	2.77	3.72	1.80
	S.D.	2.31	3.13	2.82	2.43	2.83	2.12
Females	Mean	2.33	5.98	2.69	2.77	6.19	4.18
	S.D.	2.12	3.33	2.78	2.72	3.64	3.41
t (276 d.f.)		3.22	0.74	1.18	0.01	-6.17	-7.11
Effect size (*d*)		.41	.10	.15	.00	-.79	-.91
p		$<.001$	$<.05$			$<.001$	$<.001$
d for ipsatived scores (153 of the 188 boys, 71 of the 90 girls)*		.67	.18	.52	.35	-.65	-.88

*We were never able to find the missing cases, so these *d*s are not necessarily comparable to the *d*s for the non-ipsatively-scored *HOC*.

Table 3

Gender differences on eight cognitive tests

among 188 Boys and 90 girls

		SAT-M	DAT-MR	ACT-M	ACT-NS	Alg. 1	DAT-AR	DAT-SR	SAT-V
Males	Mean	543.19	54.86	21.17	30.05	24.07*	44.40	43.84	432.87
	S.D.	69.43	7.00	6.40	7.55	6.39	3.70	9.10	72.64
Females	Mean	496.89	49.97	17.47	27.96	22.44	44.00	43.22	443.78
	S.D.	48.45	6.21	5.12	6.36	6.10	3.38	7.50	61.20
t (276 d.f.)		5.70	5.65	4.81	2.27	2.02*	0.86	0.56	-1.23
Effect size (*d*)		.73	.72	.62	.29	.26	.11	.07	-.16
p		$<.001$	$<.001$	$<.001$	$<.05$	$<.05$			

*N = 186, so 274 d.f. for the *t* test

Table 4

Correlation of the six non-ipsative Holland Occupations Checklist scores with the six Study of Values scores for 188 boys, showing corresponding categories

	Aes	The	Eco	Pol	Soc	Rel
Art	.33	-.04	-.29	-.09	.08	-.02
Inv	.04	.29	-.18	-.26	-.01	.07
Con	-.07	-.12	.23	-.03	-.03	.01
Ent	-.05	-.19	.03	.20	-.10	.06
Soc	.08	-.13	-.16	-.06	.16	.08
Rea	-.09	.05	.07	-.14	-.05	.12
	.24	-.14	-.30	-.38	.05	.32

Table 5

Correlation of the six ipsatived Holland Occupations Checklist scores with the six Study of Values scores for 188 boys, showing corresponding categories

	Aes	Eco	The	Pol	Soc	Rel
Art	.41	-.33	-.07	-.03	.12	-.10
Con	-.10	.35	-.09	.04	-.01	-.10
Inv	-.06	-.10	.32	-.18	-.03	.02
Ent	-.14	.05	-.24	.30	-.08	.08
Soc	.05	-.13	-.20	.02	.20	.05
Rea	-.17	.20	.06	-.04	-.14	.06

Table 6

Correlation of the six non-ipsative Holland Occupations Checklist scores with the six Study of Values scores for 90 girls, showing corresponding categories

	Soc	Aes	The	Eco	Rel	Pol
Soc	.42	-.04	-.26	-.17	.37	-.42
Art	.33	.39	-.36	-.27	.21	-.39
Inv	.03	.08	.18	.02	-.09	-.26
Con	.09	-.02	-.01	.18	-.04	-.17
Ent	.20	-.05	-.13	.01	.14	-.19
Rea	.25	-.06	-.06	-.02	.09	-.23
	1.32	.30	-.64	-.25	.68	-1.66

Table 7

Correlation of the six ipsatived Holland Occupations Checklist scores with the six Study of Values scores for 90 girls, showing corresponding categories

	The	Aes	Soc	Eco	Rel	Pol
Inv	.44	-.05	-.25	.06	-.30	.12
Art	-.28	.42	.07	-.26	.11	-.13
Soc	-.25	-.12	.32	-.10	.32	-.21
Con	.01	-.13	-.08	.30	-.11	.08
Ent	-.17	-.12	-.02	.12	.14	.07
Rea	.08	-.16	.04	.01	-.05	.12

Table 8

Intercorrelations of eight ability tests

(188 boys above the diagonal, 90 girls below it)

	ACT-M	SAT-M	ACT-NS	DAT-SR	DAT-AR	Alg. I	DAT-MR	SAT-V
ACT-M		.74	.31	.27	.38	.68	.11	.26
SAT-M	.56		.26	.29	.36	.63	.15	.20
ACT-NS	.35	.14		.38	.22	.18	.40	.63
DAT-SR	.38	.41	.16		.54	.11	.61	.15
DAT-AR	.34	.35	.18	.21		.20	.33	.06
Alg. I	.67	.50	.28	.29	.39		.01	.17
DAT-MR	.09	.30	-.06	.28	.16	.18		.18
SAT-V	.06	-.10	.52	-.10	.03	.02	-.14	

Table 9

Correlation of the six Study of Values scores with

eight ability-test scores for 188 boys,

ordered by row and column sums

	ACT-NS	SAT-V	DAT-SR	ACT-M	DAT-MR	SAT-M	Alg. I	DAT-AR	Row sum
The	.29[a]	.22[b]	.14	.27[a]	.13	.23[b]	.18[c]	.03	1.49
Aes	.09	.20[b]	.05	-.05	.11	.02	-.06	-.01	.35
Eco	.06	-.11	.02	.09	.06	.11	.11	.00	.34
Rel	-.09	-.08	.11	-.19[b]	-.02	-.24[a]	-.19[b]	.09	-.61
Soc	-.18[c]	-.05	-.19[b]	-.01	-.25[a]	.02	.07	-.11	-.70
Pol	-.17[c]	-.19[b]	-.21[b]	-.04	-.07	-.02	-.01	-.05	-.76
Column sum*	.88	.85	.72	.65	.64	.64	.62	.29	

*Sum of the absolute values of the rs in the column. Because of the SV's ipsativity, the algebraic sum of the rs across the SV subtests varies little from column to column (see Gleser, 1972). For rows, however--that is, within an SV category--the sum is not similarly constrained. Note, for example, that Theoretical correlates positively with each of the 8 ability-test scores, whereas Political correlates negatively with every one of them. There is no such pattern within any of the 8 columns.

a: $p < .001$ b: $p < .01$ c: $p < .05$

Table 10

Correlation of the six Study of Values scores with

eight ability-test scores for 90 girls,

ordered by row and column

	SAT-V	ACT-NS	DAT-MR	Alg.I	DAT-AR	DAT-SR	ACT-M	SAT-M	Row Sum
Aes	.23 [c]	.14	.27 [b]	.18	.14	.16	.08	-.06	1.14
The	-.22 [c]	.01	.08	.11	.05	.04	.06	.08	.21
Eco	-.13	-.07	-.03	.06	.08	.04	.05	.02	.02
Pol	-.22 [c]	-.27 [b]	-.04	-.08	.11	.00	.00	.05	-.45
Soc	-.09	-.10	-.08	-.08	-.10	.02	.02	-.09	-.50
Rel	.30 [b]	.18	-.21 [c]	-.18	-.21 [c]	-.22 [c]	-.19	-.03	-.56
Column Sum*	1.19	.77	.71	.69	.69	.48	.40	.33	

*Sum of the absolute values of the rs in the column

Table 11

Correlations between the six scales of the non-ipsative Holland Occupation Checklist and eight ability measures for 188 boys, ordered by row and column means

	DAT-MR	SAT-V	ACT-NS	SAT-M	DAT-SR	DAT-AR	ACT-M	Alg. I	Row sum
Inv	.13	.19[b]	.17[c]	.07	.08	-.06	.00	-.02	.56
Con	.12	-.01	.06	.05	.06	.01	.04	-.10	.23
Art	.00	.16[c]	-.01	.05	-.01	-.10	-.06	-.04	-.01
Rea	.16[c]	-.10	.02	-.07	.13	.03	-.11	-.13	-.07
Ent	.03	-.09	-.03	.01	-.07	.00	-.03	-.08	-.26
Soc	-.06	.01	-.07	-.06	-.20[b]	-.18[c]	-.16[c]	-.12	-.84
Column sum	.38	.16	.14	.05	-.01	-.30	-.32	-.49	-.39

Table 12

Correlations between the six scales of the non-ipsative Holland Occupations Checklist and eight ability measures for 90 girls, ordered by row and column means

	ACT-M	ACT-NS	SAT-V	DAT-MR	DAT-SR	Alg. I	SAT-M	DAT-AR	Row sum
Con	.14	.01	-.01	.11	-.03	.08	.06	.05	.41
Inv	.25 [c]	.15	.01	-.08	.11	.09	-.04	-.15	.34
Art	.14	.08	.10	-.01	.04	.00	-.08	.04	.31
Ent	-.03	.02	.09	-.08	-.09	-.17	-.13	-.15	-.54
Rea	-.04	-.06	-.10	.11	-.09	-.18	-.04	-.29 [b]	-.69
Soc	-.09	-.01	.04	-.08	-.01	-.19	-.17	-.20	-.71
Column sum	.37	.19	.13	-.03	-.07	-.37	-.40	-.70	-.88

Table 13

Correlations between the six scales of the ipsatived form of the Holland Occupations Checklist and eight ability measures for 188 boys, ordered by row and column sums

	SAT-V	DAT-MR	DAT-SR	ACT-NS	ACT-M	SAT-M	DAT-AR	Alg. I	Row sum
Inv	.18 [c]	.08	.05	.19 [b]	.06	.04	.02	.06	.68
Con	-.06	.12	.08	.06	.14	.11	.03	.02	.50
Art	.19 [b]	-.10	-.03	-.08	.00	.02	-.03	.03	.00
Rea	-.25 [a]	.11	.15 [c]	-.07	-.12	-.12	.10	-.07	-.27
Ent	-.17 [c]	-.05	-.07	-.06	-.01	-.03	.02	-.03	-.40
Soc	-.02	-.25 [a]	-.28 [a]	-.18 [c]	-.13	-.07	-.18 [c]	-.04	-1.15
Column sum*	.87	.71	.66	.64	.46	.39	.38	.25	

*Sum of the absolute values of the rs in the column

Table 14

Correlations between the six scales of the ipsatived form of the Holland Occupations Checklist and eight ability measures for 90 girls, ordered by row and column sums

	Alg. I	DAT-AR	ACT-M	DAT-MR	SAT-V	SAT-M	ACT-NS	DAT-SR	Row sum
Con	.15	.24 [c]	.14	.17	.02	.19	.02	-.05	.88
Inv	.15	-.02	.18	-.05	-.14	.02	.06	.11	.31
Art	.04	.14	-.01	-.01	.05	-.02	-.02	.00	.17
Rea	-.07	-.10	-.04	.25 [c]	-.18	.15	-.14	-.07	-.20
Ent	-.15	-.10	-.11	-.09	.25 [c]	-.12	.11	-.14	-.35
Soc	-.21 [c]	-.17	-.25 [c]	-.16	.06	-.16	-.06	.02	-.93
Column sum*	.77	.77	.73	.73	.70	.66	.41	.39	

*Sum of the absolute values of the rs in the column

Table 15

Intercorrelations of Study of Values scores for 395 bright boys and 327 bright girls tested in Iowa during 1988-1992, expressed as deviations from the mean $\underline{r}$ for the sex (-.191 for boys, -.194 for girls). Boys above the diagonal, girls below it (Lubinski, Benbow, and Sanders, 1993).

	The	Eco	Aes	Soc	Pol	Rel
The		.39	.09	-.23	.24	-.30
Eco	.37		-.11	-.30	.45	-.23
Aes	-.06	-.22		.18	-.08	-.08
Soc	-.20	-.24	.17		-.16	.37
Pol	.28	.44	-.05	-.15		-.22
Rel	-.26	-.13	.05	.24	-.25	