Journal of Counseling Psychology
1998, Vol. 45, No. 4, 436–453

0022-0167/98/$3.00

Validity of Assessing Educational–Vocational Preference Dimensions Among Intellectually Talented 13-Year-Olds

David B. Schmidt, David Lubinski, and Camilla Persson Benbow
Iowa State University

Study 1 examined the construct validity of the Strong Interest Inventory and the Study of Values for 695 intellectually talented 13-year-olds. Study 2 consisted of a generalization probe to 695 graduate students enrolled in select universities. This analysis manifested an impressive degree of adolescence-to-adult cross-validation. Well-known preference questionnaires appear to assess meaningful individual differences among intellectually talented young adolescents. How preference assessments may complement routine ability assessments of gifted adolescents and how counselors may use such information to encourage students to take a more active role in their personal development are discussed. The authors also present a methodological application, responsive to R. V. Dawis's (1992) concern about the amount of redundancy in psychological measuring tools.

For years, seventh graders who score in the top 3% on conventional standardized tests administered in their schools have been given opportunities to explore, through talent searches, whether their intellectual abilities are sufficiently precocious to warrant some form of educational intervention. Routinely, they are invited to participate in above-level ability testing, during which they take the College Board Scholastic Aptitude Test (SAT), an instrument designed for college-bound high school seniors (Cohn, 1991). Seventh graders in the top 1% of ability generate score distributions mirroring those of typical high school students (Benbow, 1988, 1990; Benbow & Stanley, 1983). Individual differences in level and pattern on the SAT hold applied psychological significance for these students (as they do for college-bound high school students). The more verbally talented tend to gravitate toward and excel in languages and the humanities, whereas the more quantitatively gifted do the same for math/science disciplines (Benbow, 1992; Benbow & Lubinski, 1996; Benbow & Stanley, 1983).

Still, abilities are only one facet of an individual's psychological makeup. Nonintellectual attributes (e.g., preferences) also need to be evaluated in order to determine correspondent learning and working environments. Traditionally, abilities and preferences have constituted the primary sources of individual differences examined in educational and vocational counseling (Dawis, 1991, 1992; Williamson, 1965). Although decades of research document the utility of ability assessments for 13-year-old gifted students with instruments initially designed for young adults, less attention has been focused on assessing their nonintellectual attributes. Could preference instruments designed for older individuals be validly extended to intellectually talented young adolescents?

Though this may seem like an early age to collect such information, a substantial body of literature suggests that talented adolescents begin thinking about career aspirations earlier than their age-mates do (Kerr & Erb, 1991; Milne, 1979; Silverman, 1993; Willings, 1986). If, for example, their preferences (i.e., interests and values) are mature enough to generate valid information, assessing these constructs might be useful in clarifying the complexity of educational, career, and development issues that this special population confronts at a younger than typical age. In this vein, Lubinski, Benbow, and Ryan (1995) provided evidence for the longitudinal stability of the Strong Interest Inventory (hereafter referred to as "the Strong"; Hansen & Campbell, 1985) over a 15-year interval (ages 13–28) for gifted adolescents; Lubinski, Schmidt, and Benbow (1996), using an independent sample, documented the longitudinal stability of Allport, Vernon, and Lindzey's (1970) Study of Values (SOV) over a 20-year interval (ages 13–33). Furthermore, gifted young adolescents have displayed a wide range of individual differences on both the Strong and the SOV (Achter, Lubinski, & Benbow, 1996), such that further examination of these measures, and the constructs they assess, appears promising.

Yet, a promising degree of temporal stability, even when accompanied by other within-instrument statistical properties routinely observed in adult populations, is insufficient to justify extending questionnaires initially designed for much older people to this special population in ways currently enjoyed by ability tests (e.g., the SAT). One more step needs to be taken. Before practitioners are encouraged to consider implementing these tools in practice, external validation is needed. We need to know whether preference questionnaires

David B. Schmidt, David Lubinski, and Camilla Persson Benbow, Department of Psychology, Iowa State University.

This article is based on a master's thesis completed by David B. Schmidt under the supervision of David Lubinski and Camilla Persson Benbow. Part of this study was supported by the Strong Research Advisory Board, Consulting Psychologists Press, and an anonymous service agency.

Correspondence concerning this article should be addressed to David B. Schmidt, who is now at Aon Consulting, P.O. Box 36778, Grosse, Michigan 48236 or David Lubinski, who is now at Department of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee 37203. Electronic mail may be sent to David_Schmidt@aoncons.com or David.Lubinski@Vanderbilt.edu.

administered to intellectually gifted adolescents provide the same kind of information they routinely provide the adult populations for which they were designed. If they do, such assessments might provide valuable information for educators and psychologists working, for example, in designing educational programming or in facilitating educational/vocational counseling and development. Support for these ideas may be gleaned from developmental psychology in the work of Sandra Scarr.

According to Scarr (1996), valid information about relatively stable personal attributes (abilities, personality, and preferences) can be useful for refining psychological practice by providing students with important information. This information would allow them to take a more active role in designing environments for themselves ("niches"), which are likely to engender more personally meaningful learning experiences (Lubinski, 1996). This underscores the emphasis Dawis (1992) placed on multifaceted assessment and the importance of a multidimensional approach. It also forestalls certain problems encountered in research on gifted students emanating from assessing a single personal-attribute domain in isolation (e.g., exclusively focusing on abilities). Such restricted assessments frequently result in "misdiagnoses" and attributions of "multipotentiality" (flat profile). When one assesses multiple domains with developmentally appropriate instruments, however, this conclusion is found infrequently. For example, Achter et al. (1996) found that multipotentiality (genuinely flat profiles) occurred in less than 5% of the gifted population when abilities, interests, and values were considered conjointly. This should not be overly surprising, inasmuch as individuals with uniformly high verbal, mathematical, and spatial abilities typically manifest substantial individual differences in preferences. Imagine, for example, that two individuals have high-flat ability profiles, but one desires a great deal of social stimulation, while the other desires little. Certainly, each is likely to prefer somewhat different educational and vocational contexts and would likely seek different environments in which to express their abilities as well as their differential degrees of gregariousness in the world of work.

All of this naturally follows from traditional models of person–environment fit, which is one reason the theory of work adjustment (TWA; Dawis & Lofquist, 1984; Lofquist & Dawis, 1991) undergirds our research. This model devotes equal attention to person and environment factors for maximizing adjustment. We have used TWA principles to design learning environments for gifted youth (Lubinski & Benbow, 1995). According to TWA, optimal learning and work environments are defined by two broad dimensions of correspondence: satisfactoriness (the match between an individual's abilities and the ability requirements of the environment) and satisfaction (the match between an individual's preferences and the rewards offered by the environment). To the extent that both satisfactoriness and satisfaction are achieved, both the individual and environment are motivated to continue to interact with one another—to change, learn, and profit from each another. Correspondence on both dimensions is likely to lead to a long-term, profitable relationship between the individual and environment. With discorrespondences on either dimension, the relationship is less likely to continue for an extended period of time. Returning to our example of the two students with different social contact needs, if both were placed in an environment with high social demands, it is unlikely that both would thrive equally well. These aspects of individuality need to be factored into applied psychological frameworks.

Talent searches using concepts from TWA have shown that ability tests designed for older individuals (e.g., SAT) are useful for estimating what gifted adolescents are capable of achieving (forecasting learning and work performance) and, thus, predicting satisfactoriness (Benbow & Lubinski, 1996). In the presented research, we asked whether preference questionnaires designed for older individuals provide valid information for estimating what gifted adolescents are likely to find motivating (personally fulfilling in learning and work) and predicting satisfaction. The applied use of models like TWA is, however, predicated on the construct validity of relevant instrumentation. We chose to examine the Strong (Hansen & Campbell, 1985; Harmon, Hansen, Borgen, & Hammer, 1994) and Allport et al.'s (1970) SOV—two of the most widely used instruments in their respective domains. Moreover, they have different dimensional features, suggesting that they are getting at somewhat distinct aspects of individuality (Lubinski et al., 1996).

A final focus of the current research stemmed from a discussion by Dawis (1992) in his *Journal of Counseling Psychology (JCP)* centennial feature. Specifically, Dawis expressed concern about the amount of redundancy in the multitude of instruments currently used by counseling psychologists. For example, in one *JCP* volume alone, Dawis counted *115* instruments that were either new or little known. In reaction to this, he stated,

> On the one hand, one cannot help applauding this impressive display of creative activity. On the other hand, one has to wonder how much of the effort is overlapping and redundant. Only occasionally does someone . . . attempt to assess the overlap among measures, with illuminating results, but such studies are few because they demand too much time from research participants. One is left only with the optimist's anticipation that things tend to sort themselves out and that in the end the cream rises to the top. (p. 16)

The magnitude of redundancy found in psychological assessment operations is not a problem restricted to counseling psychology. It has been stressed in other psychological contexts (Lubinski & Dawis, 1992; Sanders, Lubinski, & Benbow, 1995; Tellegen, 1993), and it makes for an awfully cluttered collection of tools (and labels) in our discipline. In the present research, we addressed the amount of redundancy running through interests and values as measured by the Strong and the SOV.

Given the concern over the number of instruments (and their rate of development), it seemed worthwhile to examine the extent to which the Strong and SOV measure unique sources of individual differences. These two instruments are ideal to examine this question because they epitomize the interest/value distinctions outlined by Dawis (1991). For example, the Strong requires individuals to indicate their like or dislike of or indifference to various academic

subjects, hobbies, people, work, and occupations; the SOV requires individuals to assign varying weights to preferences for different choices based on what they feel is important. The Strong is normatively scaled and provides interindividual feedback relative to a reference group; the forced-choice format of the SOV produces ipsative scaling (i.e., the scores on the six scales always add up to the same number) and provides intraindividual feedback relative to the individual. In other words, one can express as many likes or dislikes as one wants on the Strong, but the SOV forces one to decide between choices. Thus, these instruments constitute prototypical tools for assessing their respective domains and are ideal for investigating this concern. Is it possible, however, that these two instruments are simply providing redundant information?

Early research on the overlap of these two instruments indicated an appreciable amount of redundancy, but each appeared to manifest some specificity of psychological importance (Duffy & Chrissy, 1940; Ferguson, Humphreys, & Strong, 1941; Sarbin & Berdie, 1940). These early studies focused on the six SOV themes and their relationship with a subset of the Strong's empirically keyed Occupational scales. Both instruments, however, have changed substantially since the early 1940s. For example, the Strong now includes General Occupational Themes based on Holland's (1985) six theoretical dimensions and a set of theoretically developed Basic Interest scales. Although both instruments have changed substantially, our literature review failed to find empirical work examining the updated Strong and SOV conjointly, especially in terms of their incremental validity relative to each other. Two studies examining the overlap of the six SOV themes and Holland's six dimensions found evidence for redundancy, but each seemed to display an appreciable amount of uniqueness as well (Laudeman & Griffeth, 1978; Williams, 1972). None of the aforementioned studies directly assessed the redundancy of the modern Strong and SOV. Their vast developmental differences notwithstanding, researchers still do not know whether each instrument provides unique and psychologically meaningful information relative to the other. So it is certainly reasonable to speculate on the amount of overlap running through each instrument, especially given the thematic overlap among some of their scale labels' conceptual definitions (e.g., the Strong Social and SOV Social, Strong Artistic and SOV Aesthetic, and Strong Investigative and SOV Theoretical dimensions).

In the current research, based on a large sample of intellectually gifted young adolescents (Study 1) and a large sample of graduate students (Study 2), we took the equivocal overlap between contemporary versions of the Strong and SOV into account and evaluated (a) the construct validity of the Strong and the SOV, as well as their incremental validity relative to each other across ability, biographical, and personality measures; (b) whether it is necessary to administer both the Strong and the SOV for applied purposes; and (c) whether the internal and external preference structure among gifted adolescents, as assessed by the Strong and SOV, generalizes to highly talented graduate students.

Study 1

In Study 1 we examined the construct validity of the Strong and the SOV for intellectually gifted adolescents and evaluated whether the two instruments assess unique and psychologically meaningful individual differences relative to one another. Results are grouped under, but not restricted to, four validation analyses: construct validity, incremental validity, cross-validation, and extrinsic convergent validity.

Method

Participants

Participants for this study (416 male adolescents and 279 female adolescents; 88.1% Caucasian, 8.6% Asian American, .8% African American, .3% Hispanic, 1.1% other, and 1.1% not responding) were identified by the Study of Mathematically Precocious Youth's (SMPY's) 1992, 1993, and 1994 talent searches, conducted by the Office of Precolleagiate Programs for the Talented and Gifted (OPPTAG) at Iowa State University. Students who scored in the top 3% on standardized achievement tests administered in their schools were administered the SAT.[1] Students who scored 390 or higher on the Math SAT or 370 or higher on the Verbal SAT before age 13 and attended OPPTAG's summer programs for the gifted were included, as were students who scored 20 or higher on one American College Test subtest.[2] These students represent approximately the top 1% in intellectual ability (Lubinski & Benbow, 1994). Only students who completed both the SOV and the Strong were included in this study.

Instruments

The Strong (Hansen & Campbell, 1985). The Strong contains 23 Basic Interest scales and 6 General Occupational Themes, based on Holland's (1985) RIASEC dimensions: *R*ealistic (interest in working with things or working outdoors and need for structure), *I*nvestigative (interest in sciences, particularly mathematics and physical sciences, and a preference to work independently), *A*rtistic (interest in writing, art, or other creative expression and little need for structure), *S*ocial (interest in people and in helping professions), *E*nterprising (interest in leadership roles, especially if they lead to achieving economic goals), and *C*onventional (preference for structured environments, a well-defined chain of command, and office practices). Holland proposed a hexagonal structure to represent the interrelationships of these six types, such that types that were more similar were arranged physically closer than were types that were less similar. Substantial support for the hexagonal organization of the RIASEC over other proposed structures has been found in adult samples (Rounds & Tracey, 1993; Tracey &

[1] In March 1994, the SAT underwent some structural changes. The portion with Verbal and Math sections is now referred to as the SAT-I: Reasoning Test. The Verbal section now includes more reading passages and testing vocabulary in context. The Math section has more emphasis on application, interpretation, and student-produced answers. For the current study, most students were identified using the older SAT. Those taking the newer SAT were selected using cutoff scores at an ability level comparable to that of other students identified.

[2] Recently the American College Test (Math, Science Reasoning, English, and Reading subtests) also has been used to identify gifted students. Scores on subtests range from 1 to 36.

Rounds, 1993), as well as in intellectually gifted adolescent samples (Lubinski et al., 1995). The Basic Interest scales of the Strong reflect components of the more general RIASEC themes. These scales, preceded by the RIASEC letter with which they are associated, are as follows: *R* (Agriculture, Nature, Adventure, Military Activities, and Mechanical Activities), *I* (Science, Mathematics, Medical Science, and Medical Service), *A* (Music/Dramatics, Art, and Writing), *S* (Teaching, Social Service, Athletics, Domestic Arts, and Religious Activities), *E* (Public Speaking, Law/Politics, Merchandising, Sales, and Business Management), and *C* (Office Practices).

The SOV (Allport et al., 1970). The SOV assesses the relative prominence (i.e., intraindividually) of personality-related values based on Spranger's (1928) six theoretical types: Theoretical (values discovery of truth, and interests are empirical, critical, and rational), Economic (values that which is useful and practical and sees unapplied knowledge as wasteful), Political (values power and desires personal power, influence, and renown), Aesthetic (values form and harmony and is interested in the artistic side of life), Social (values altruistic/philanthropic love of others and is unselfish and sympathetic), and Religious (values unity, and tries to comprehend the cosmos and relate it to the self).[3]

Validation Criteria

Adjective Check List (ACL; Gough & Heilbrun, 1983). This instrument is designed to assess personality attributes. It contains 300 items and 37 scales, of which the following 20 were used in Study 1 (descriptions of high scorers follow in parentheses): Dominance (seeks to control relationships and seeks/maintains leadership roles), Endurance (persists in tasks undertaken), Order (emphasizes neatness, organization, and planning), Intraception (tries to understand behavior of self and others), Nurturance (engages in behaviors that provide benefits for others), Affiliation (seeks and maintains many personal friendships), Autonomy (acts independently of others or of social values/expectations), Aggression (engages in behaviors that harm others), Change (seeks novel experiences and avoids routine), Succorance (solicits sympathy, affection, or emotional support from others), Deference (seeks/maintains subordinate roles in relationships), Self-Control (is cautious, overcontrolled, conservative, patient, and quiet), Self-Confidence (is confident that goals will be achieved and is determined, assertive, and enterprising), Creative Personality (is clever, original, artistic, versatile, and imaginative), Critical Parent (is bossy, demanding, impatient, and suspicious), Nurturant Parent (is forgiving, appreciative, helpful, loyal, and stable), High Origence/Low Intellectance (has strong instincts, enjoys festivity and is easily distracted), High Origence/High Intellectance (is indifferent to convention, has original thoughts and perceptions, is aesthetically sensitive and insightful), Low Origence/Low Intellectance (is conventional, easygoing, and forthright; respects rules; and is content with life role), and Low Origence/High Intellectance (is analytical, logical, intellectual, and self-disciplined). For these 20 scales, Gough and Heilbrun (1983) reported an internal consistency median of .78 for men (interquartile range: $Q_1 = .69$, $Q_3 = .81$) and a median of .77 for women ($Q_1 = .69$, $Q_3 = .78$) and a 6-month test–retest reliability median of .68 for men ($Q_1 = .59$, $Q_3 = .74$) and a one-year test–retest median of .71 ($Q_1 = .64$, $Q_3 = .77$) for women. The median one-year test–retest reliability for SMPY participants ($N = 203$) on these 20 scales was .65 ($Q_1 = .62$, $Q_3 = .67$).

Multidimensional Personality Questionnaire (MPQ; Tellegen, 1982; Tellegen & Waller, in press). The MPQ assesses major distinctive personality traits across a broad domain. It consists of 300 items that form 11 primary scales (descriptions of high scorers follow in parentheses): Well-Being (has a cheerful, happy, and optimistic disposition), Social Potency (is forceful, decisive, and persuasive and enjoys leadership roles), Achievement (enjoys difficult/demanding tasks; works hard; and is persistent, ambitious, and perfectionistic), Social Closeness (is warm, affectionate, and sociable), Stress Reaction (is tense, nervous, worrisome, and easily upset), Alienation (feels deceived, betrayed, and used and believes others wish him or her harm), Aggression (enjoys upsetting or frightening others and is physically aggressive and vindictive), Control versus Impulsivity (is reflective, cautious, plodding, rational, and detail oriented), Harm Avoidance (dislikes danger, risk, adventure, and sudden emergencies), Traditionalism (endorses religious institutions, follows strict child-rearing practices, and has high moral standards), and Absorption (is enraptured by stimuli, thinks in images, has vivid experiences, and is immersed in own thoughts). Tellegen and Waller (in press) reported an internal consistency median of .85 ($Q_1 = .83$, $Q_3 = .88$) and a 30-day test–retest reliability median of .89 ($Q_1 = .82$, $Q_3 = .90$). The median one-year test–retest reliability for SMPY participants ($N = 201$) was .67 ($Q_1 = .61$, $Q_3 = .71$).

Family Environment Scale (FES; Moos & Moos, 1986). The FES assesses the perception of social and environmental characteristics of families. There are 90 items forming 10 scales, of which the following 3 were used in Study 1: Achievement Orientation (activities are achievement oriented and in a competitive framework), Intellectual–Cultural Orientation (interest in political, intellectual, and cultural activities), and Moral–Religious Emphasis (emphasis on ethical and religious issues and values). Moos and Moos (1986) reported internal consistencies of .64, .78, and .78 and one-year test–retest reliabilities of .69, .79, and .89, for Achievement Orientation, Intellectual–Cultural Orientation, and Moral–Religious Emphasis, respectively. One-year test–retest reliabilities for SMPY participants ($N = 245$) were .66, .64, and .80, respectively.

Mental Rotation Test (MRT; Vandenberg & Kuse, 1978). The MRT measures three-dimensional spatial visualization with a paper-and-pencil test. It contains two timed 5-min sections with 10 items each. Participants are required to match a criterion figure to two identical, yet rotated, figures out of four possible figures. Points are awarded for identifying one or two identical figures, but no points are awarded if any part of an answer is incorrect. Vandenberg and Kuse (1978) reported an internal consistency of .88, and Kuse (1977) reported one-year test–retest reliabilities on two samples as .83 and .70. One-year test–retest reliability for SMPY participants ($N = 255$) was .73.

Mechanical Comprehension Test (MCT) Form S (Bennett, 1969, 1994). This instrument assesses grasp of physical and mechanical relationships in practical situations. The MCT is a timed 30-min test containing 68 multiple-choice items. Bennett (1969, 1994) reported internal (split-half) consistencies ranging from .81 to .93 (median = .86). One-year test–retest reliability for SMPY participants ($N = 255$) was .85.

Advanced Progressive Matrices (Raven, Court, & Raven, 1977). This 36-item, untimed test assesses nonverbal reasoning ability. Each item requires examining a pattern of figures and determining which of several options completes the pattern. Raven (1973) reported 2-month test–retest reliabilities of .91 for adults, .86 for 12-year-olds, and .76 for 10-year-olds. One-year test–retest reliability for SMPY participants ($N = 255$) was .60.

[3] The SOV was modernized by SMPY researchers to contain gender-neutral descriptors. Item changes (primarily pronoun changes) were easily made and seemingly without altering their psychological meaning in any substantively significant way.

Background questionnaire. This questionnaire was designed by SMPY and asks questions related to attitudes, demographics, family background, and future plans, with a particular emphasis on educational and vocational intentions. In the current study, the following items were included: reading preferences ("How many books/magazines have you read in the last 12 months, not including those required for school, in the following categories?" Participants responded to Science, Love Stories, Plays/Poetry/Essays, and Religious), future occupational importance of fields of study ("When you think about your future occupation, how important do you think skills in the following domains will be?" Participants responded to Math, Physics, Chemistry, Biology, Reading, Writing, Social Studies, Foreign Language, and Computer Science), ratings of academic subjects ("List your three favorite courses in school"), lifestyle ratings ("How important is lots of money in your life? Strong friendships? Being a community leader?"), and some autobiographical questions about hobbies ("To what extent were you involved with 'tinkering' with equipment, mechanical gadgets, or construction games as a young child? As an adolescent?").

Procedure

Intellectually able participants came to summer programs sponsored by OPPTAG to take high school or college courses at an accelerated pace. As part of their experience, students completed tests and questionnaires for SMPY's ongoing longitudinal study for facilitating their educational and vocational decision-making process. Questionnaires were completed at home; ability tests were administered in a 2–2.5 hr standardized mass-testing session (Lubinski & Benbow, 1994).

Results

Means and standard deviations of the SOV and the Strong (i.e., RIASEC and Basic Interest Scales) are provided by gender and are accompanied by male/female effect sizes in Table 1. On the SOV, boys scored highest on Theoretical and lowest on Religious, while girls scored highest on Aesthetic (followed closely by Social and Theoretical) and lowest on Religious. Consistent with previous SOV literature (Allport et al., 1970), boys tended to score higher than girls on Theoretical–Economic–Political values, whereas girls tended to score higher on Aesthetic–Social–Religious values. Among the RIASEC dimensions, boys scored highest on Investigative and lowest on Social, while females scored highest on Artistic (followed closely by Investigative) and lowest on Realistic (followed closely by Enterprising). Examination of male/female effect sizes on Basic Interest scales revealed that girls tended to score higher than boys on 15 of 22 Basic Interests scales. Further, the girls' average exceeded 50 on 9 Basic Interests scales, as compared to 5 for boys. This is not surprising, given earlier findings that intellectually talented girls tend to have more eclectic interests (and values), and boys tend to be more differentiated (Achter et al., 1996; Lubinski et al., 1995; Lubinski et al., 1996).[4]

Construct Validity

To evaluate the construct validity of the Strong and the SOV, we examined their intercorrelations and external relationships across 59 relevant external criteria. Table 1 presents correlations between scales on the Strong and SOV.[5] For the most part, the SOV and Strong scales that were expected to covary did. For example, Theoretical values covaried positively with Investigative and negatively with Artistic and Social interests. Political values covaried positively with Athletics, Adventure, Law/Politics, and negatively with Artistic interests. Aesthetic values covaried positively with Artistic interests. Social values covaried positively with Social (especially Social Service) and Artistic interests. Finally, Religious values covaried highest with interests in Religious Activities (.70).

Next, we considered relationships between the SOV and RIASEC themes and external criteria presented in Table 2. Following Cattell's (1965) recommendation, we used criteria from three chief data sources: life record and biographical information (L data), subjective questionnaires (Q data), and objective tests (T data). Moreover, following Fiske's (1971) recommendation for conducting extrinsic convergent validation analyses (see below), we examined validation criteria all along the convergent–discriminant continuum. As shown in Table 2, we found a convergent and discriminant pattern that was in good accord with what one might expect on the basis of well-known educational/vocational preference findings secured on older populations (Dawis, 1991; Harmon et al., 1994).

Though we do not detail all patterns in Table 2, a few examples are in order. Theoretical values covaried positively with the future occupational importance of math and physics, tinkering behaviors, preference for science courses, and mechanical and spatial ability and negatively with FES Moral–Religious Emphasis, MPQ Social Closeness, ACL Nurturance, and preference for humanities courses. Religious values covaried positively with MPQ Traditionalism, FES Moral–Religious Emphasis, and reading of religious books and negatively with the importance of money. Investigative interests covaried positively with MPQ Achievement; future occupational importance of math, biology, chemistry, and physics; preference for science courses; and mathematical ability. Social interests covaried positively with ACL Nurturance, Affiliation, and Nurturant Parent and also with

---

[4] Because the remaining analyses dealt primarily with the structural properties of the Strong and the SOV, we combined male and female data. Following Huba and Hamilton's (1976) suggested method for quantifying the degree of isomorphism between intercorrelation matrices, corresponding male/female intercorrelations for the Strong and the SOV were themselves correlated and manifested values of .89 and .97, respectively. This suggested that despite male/female mean differences, the covariances were quite similar and thus could be combined for structural analyses.

[5] When examining SOV validity coefficients, it is important to keep in mind the constraints placed on the SOV scales as a result of ipsativity (Clemans, 1966; Hicks, 1970). Ipsativity forces a negative manifold for within-instrument intercorrelations (the estimated average interscale correlation is −.20). This property attenuates observed correlations between the SOV and other measures (e.g., RIASEC dimensions) on constructs conceived of as theoretically related (e.g., Investigative/Theoretical, Artistic/Aesthetic, Social/Social, and Enterprising/Economic). Thus it is especially important to take into consideration the pattern of relationships between the SOV and construct validation criteria, rather than exclusively focusing on the magnitude of the validity coefficients.

Table 1
Correlations Between Study of Values (SOV) and Strong Interest Inventory Scales, Along With Means and Standard Deviations by Gender

Strong scale	SOV scale: Theoretical	Economic	Political	Aesthetic	Social	Religious	Boys *M*	Boys *SD*	Girls *M*	Girls *SD*	Effect size (boys − girls)
RIASEC dimension											
Realistic	.15	.13	.09	−.16	−.17	−.02	48.3	9.2	43.1	7.8	.60
Investigative	.29	−.06	−.05	−.06	−.05	−.06	53.2	8.2	52.4	9.0	.09
Artistic	−.33	−.45	−.29	.53	.29	.12	43.0	9.9	53.2	9.7	−1.04
Social	−.39	−.34	−.16	.03	.42	.30	39.7	10.1	48.7	10.2	−.89
Enterprising	−.20	.06	.16	−.06	.07	.00	43.1	9.5	44.4	8.9	−.14
Conventional	.01	.18	.11	−.20	.00	−.05	47.1	9.7	46.5	9.6	.06
Basic Interest scale											
Agriculture	−.12	−.06	.00	−.04	.05	.13	43.6	8.2	44.4	7.8	−.10
Nature	−.17	−.26	−.22	.20	.27	.10	41.5	9.4	49.1	9.8	−.79
Adventure	.09	.07	.25	−.10	−.18	−.07	56.3	9.4	51.3	10.0	.52
Military Activities	.05	.05	.21	−.16	−.14	.02	50.9	10.1	46.6	8.8	.45
Mechanical Activities	.31	.21	.06	−.21	−.24	−.08	50.8	10.0	45.3	8.2	.59
Science	.47	.07	−.06	−.20	−.16	−.08	57.9	7.3	54.4	8.7	.44
Math	.34	.17	.01	−.24	−.18	−.06	57.7	7.6	54.0	9.7	.44
Medical Science	.04	−.05	.02	−.06	.08	−.03	47.5	9.7	49.4	9.4	−.20
Medical Service	−.05	−.04	−.05	−.03	.16	.00	48.0	7.9	51.1	9.8	−.36
Music/Dramatics	−.40	−.47	−.34	.48	.34	.22	43.2	8.6	54.8	9.5	−1.29
Art	−.31	−.41	−.30	.56	.26	.09	43.1	9.5	53.4	10.1	−1.06
Writing	−.28	−.47	−.23	.41	.32	.12	43.2	9.6	53.0	9.5	−1.03
Teaching	−.27	−.36	−.23	.18	.38	.18	43.4	10.0	52.2	10.2	−.87
Social Service	−.42	−.42	−.21	.20	.53	.19	40.2	8.6	50.7	10.0	−1.14
Athletics	−.11	.15	.34	−.31	−.04	.03	49.4	9.7	45.4	8.9	.43
Religious Activities	−.39	−.37	−.30	−.15	.22	.70	42.9	9.5	46.7	9.9	−.39
Public Speaking	−.24	−.26	.21	−.06	.18	.15	44.5	9.8	47.3	10.0	−.28
Law/Politics	−.14	−.13	.33	−.08	.07	.00	46.4	9.6	48.0	9.8	−.17
Merchandising	−.31	−.02	.03	.09	.18	.02	40.6	8.8	45.1	9.3	−.50
Sales	−.17	.12	.22	−.07	.01	−.05	46.4	8.3	45.5	7.8	.11
Business Management	−.18	.10	.16	−.10	.05	.00	40.9	9.4	42.7	8.9	−.20
Office Practices	−.17	.01	−.05	−.01	.16	.03	45.7	7.5	48.7	9.3	−.36
Boys											
M	47.6	43.0	43.4	37.6	35.8	32.5					
SD	7.1	7.6	6.5	7.5	7.4	10.9					
Girls											
M	41.7	38.0	39.2	43.8	41.8	35.4					
SD	7.9	6.7	6.6	8.3	7.4	10.0					
Effect size (boys − girls)	.79	.69	.64	−.79	−.81	−.27					

Note. Male $n = 416$, and female $n = 279$. For total sample, correlations of .10 or higher are statistically significant at $p < .01$. For male–female effect sizes greater than .20, $p < .01$. Strong = Strong Interest Inventory.

MPQ Well-Being and Social Closeness. Social interest covaried negatively with ACL Autonomy, Aggression, and Critical Parent and with MPQ Aggression and mechanical/spatial ability. As can be seen from these few examples, the patterns of correlates suggest that these scales covaried with external criteria as one might expect. The Basic Interest scales also were examined and, across the 59 external criteria, followed theoretical expectations gleaned from the adult literature (Harmon et al., 1994).[6]

Some clusterings of abilities (verbal, mathematical, spatial, and mechanical reasoning) and preferences (RIASEC and SOV) are worth emphasizing because they are in line with Ackerman's (1996; Ackerman & Heggestad, 1997) findings with adult samples. For example, Theoretical and Realistic themes displayed marked positive correlations with spatial ability (.25 and .20, respectively) and mechanical reasoning (.26 and .23, respectively), whereas SOV Social and Strong Social and Artistic displayed negative correlations with spatial ability (−.18, −.21, and −.17, respectively) and mechanical reasoning (−.23, −.28, and −.20, respectively). The largest Math SAT correlation

[6] To conserve space, we do not present correlations between the Strong's Basic Interest scales and external criteria. Some of the larger correlations were Medical Science with anticipated occupational importance of biology (.52) and chemistry (.39); Science and Math with anticipated occupational importance of math (.36 and .46, respectively) and physics (.43 and .39, respectively); Mechani-

across these 12 measures was with Investigative (.27), whereas for Verbal SAT it was Artistic (.29), followed by Aesthetic (.21). Throughout Table 2, the correlational patterning of these general themes depicts a humanistic versus scientific orientation to study and work. Support also was found in convergent and discriminant correlations observed with the Strong's Basic Interest scales: Math and Math SAT = .39, versus Math and Verbal SAT = .00; Writing and Verbal SAT = .37, versus Writing and Math SAT = .04. Mechanical Activities correlated .32 and .28 with mechanical reasoning and spatial visualization, respectively, and Mechanical Activities correlated .03 and −.12 with Math and Verbal SATs, respectively.

Incremental Validity

The correlations of the SOV and Strong with each other and with external criteria provided support for their construct validity for gifted youth. To specify the applied utility of these instruments more precisely, we wanted to determine the relationship between the external criteria we used and the SOV and Strong as entire instruments. That is, how much information, or total variance, can these two instruments account for, separately and together, with respect to the criteria we examined? To answer this question, we computed multiple correlations between each criterion variable and the two instruments separately and then together. We first computed the multiple correlation between each external criterion and five SOV themes as a block.[7] We next computed the multiple correlation between each external criteria and 28 Strong (RIASEC and 22 of the 23 Basic Interest) scales as a block. We excluded Domestic Arts because it is not included on the revised 1994 Strong (it is worth noting that this scale made no significant contribution to the multiple correlations). Finally, we computed the multiple correlation between each external criterion and all 33 SOV and Strong scales as a block.

The multiple correlations of each external criterion variable with each block of scales are presented in Table 2. For all criterion variables, the Strong displayed comparable or greater multiple correlations with the external criteria compared with the SOV. Moreover, with only a few exceptional cases (MPQ Traditionalism, ACL Nurturance, and FES Moral–Religious Emphasis), the SOV appeared to add little incremental validity to the Strong, at least relative to the 59 criteria examined here. Examining the corresponding correlations in the last three columns of Table 2 tells the story. For each block of scales, the average multiple correlation with the external criteria was as follows: for the SOV, R = .28; for the Strong, R = .46; and for the SOV and the Strong, R = .48.

Cross-Validation

The above findings provide encouraging evidence regarding the construct validity of the Strong and the SOV for intellectually gifted young adolescents. It also is clear that the SOV adds little incremental validity beyond the Strong with respect to these criteria. This motivated us to analyze the extent to which the SOV could be reproduced via Strong-based multiple regression equations. This is an especially relevant question for applied purposes, because, if the Strong truly does absorb the SOV, it certainly would not be necessary to administer both instruments in all settings. How much of the actual SOV profile is obtainable from multiple regression equations using the Strong's 28 theoretical dimensions? Because of the particulars described below, our sample of 695 afforded us the opportunity to address this question in a number of ways.

Two subsamples. OPPTAG students tend to enjoy their summer school experience (Benbow, Lubinski, & Suchy, 1996), and many return the following summer for further learning opportunities. Among the 695 participants currently under analysis, 110 participated in consecutive summers and took the Strong and SOV (1992–1993 or 1993–1994). Therefore, we split our sample into two independent subsamples: We used the 585 nonrepeaters to develop Strong-based multiple regression equations to predict each of the six SOV scales; the remaining 110 repeaters were used as the cross-validation sample for these equations. Following Lord and Novick's (1968) nomenclature for conducting cross-validation analyses, the nonrepeaters constitute a *screening* (developmental) sample, whereas the repeaters constitute a *calibration* (cross-validation) sample. Because the members of the calibration sample completed the Strong and SOV at two points in time, several cross-validations (i.e., concurrent, predictive, and retrospective) were possible. In addition, SOV test–retest stabilities could serve as baselines for comparing cross-validation coefficients for each scale. The particulars of our cross-validation analyses are illustrated in Figure 1. We first developed multiple regression equations,

cal Activities with current tinkering (.51); Adventure with MPQ Harm Avoidance (−.54); Music, Art, and Writing with MPQ Absorption (.36, .35, and .36, respectively); Public Speaking and Law/Politics with MPQ Social Potency (.46 and .42, respectively); and Religious Activities with FES Moral–Religious Emphasis (.50) and MPQ Traditionalism (.41). We also noted some interesting patterns. Although Science, Math, Medical Science, and Medical Service all covaried positively with future occupational importance of math and the physical sciences, Medical Science and Medical Service covaried much higher with the anticipated occupational importance of biology and chemistry, whereas Science and Math covaried much higher with the anticipated occupational importance of math and physics; furthermore, Science and Math covaried much higher with preferences for science courses and tinkering activities. A second pattern that emerged was that Music/Dramatics, Art, Writing, Teaching, and Social Service covaried positively with future occupational importance of reading, writing, social studies, and foreign languages. These scales also covaried positively with preferences for humanities courses and negatively with preferences for science courses and tinkering activities.

[7] We used only five scales for the SOV, because the inventory's ipsative nature means that all the information on any sixth scale may be derived from the remaining five (SOV profiles always sum to 240). Omitting any one of the six scales results in uniform multiple correlations for the remaining five.

using the 28 Strong scales with the screening sample, to forecast each of the SOV themes.[8] Next, we applied the six regression equations to the 110 participants in the calibration sample at both time points. This gave us two SOV forecasts for each SOV theme (one from Time 1 Strong, and the other from Time 2 Strong). Following Figure 1, we examined the relationships between these two forecasted SOVs and the two actual SOVs for cross-validation purposes. Furthermore, the one-year test–retest stabilities of the two forecasted SOVs and the two actual SOVs were examined.

Interindividual analyses. One-year test–retest coefficients for the six actual-SOV themes are presented in the first column of Table 3. These served as a baseline for comparing the test–retest coefficients for the six forecasted-SOV themes, as well as concurrent, predictive, and retrospective cross-validation coefficients (see Table 3). The actual SOVs were relatively stable, but overall the forecasted SOVs were a bit more stable. Across the board, both within and between time-frames, cross-validation coefficients were comparable to actual-SOV stability coefficients. The coefficients were essentially equivalent, regardless of whether we were correlating an actual with a forecasted SOV, two forecasted SOVs, or two actual SOVs. Remarkably, this generalization pertains to forecasts generated for concurrent as well as predictive assessments having a one-year time lag. These analyses provide another piece of evidence that the SOV can be captured with regression equations based on the Strong.

To examine the possibility that the sheer number of Strong scales used resulted in high cross-validation coefficients, we developed equations using the best six Strong scales for predicting each SOV theme. Comparing the cross-validation coefficients resulting from these equations with those using all 28 Strong scales (see Table 3), we found little difference in the magnitude of the coefficients. We subsequently used only the equations based on all 28 Strong scales, because all scales are readily available on the Strong and regression procedures that retain a selected set of variables from a total set are more prone to capitalizing on chance factors (Pedhazur, 1982).

Intraindividual analyses. We next examined the rank order of the six SOV themes in the actual and forecasted profiles. Many educational/vocational counselors using an instrument such as the SOV tend to advise clients in accordance with their most salient themes. Actually, for many counseling issues, intraindividual profiling is central, because it readily enables counselors to focus on client strengths and areas of chief importance (Dawis, 1992). Thus, it is critical to assess the rank order of scales for actual and forecasted SOVs intraindividually. The interchangeability of actual and forecasted profiles was evaluated by computing Pearson product–moment correlations and Spearman rank-order correlations intraindividually for all 110 participants. This produced a distribution of 110 coefficients whose central tendency and dispersion (Table 4), quantify the consistency of the SOV profiles (i.e., stability of rank orders). Once again, although the actual SOVs were relatively stable over 1 year, the forecasted SOVs were more stable. The cross-validation analyses indicated that the relationship between the actual and forecasted SOVs is slightly more consistent than the two actual SOVs, except in the case of the retrospective cross-validation. As in the previous analyses, the results were similar whether we compared two actual SOVs, two forecasted SOVs, or an actual and a forecasted SOV. In other words, the same general rank order of themes was attained regardless of whether Strong-based regression equations were used to generate a SOV profile or the actual SOV was given.

---

[8] All variables in the regression equations followed the same order presented as the first (with the *y*-intercept as the first bracketed value). Theoretical = [43.91 + (Realistic × −.13) + (Investigative × .13) + (Artistic × −.23) + (Social × .03) + (Enterprising × .05) + (Conventional × −.03) + (Agriculture × −.06) + (Nature × −.02) + (Adventure × .00) + (Military Act × .08) + (Mechanical Act × .13) + (Science × .36) + (Math × .01) + (Medical Sci × −.04) + (Medical Serv × .02) + (Music × −.07) + (Art × .11) + (Writing × .10) + (Teaching × .01) + (Social Serv × −.10) + (Athletics × −.07) + (Religious Act × −.23) + (Public Speaking × .03) + (Law/Politics × −.07) + (Merchandis × −.13) + (Sales × −.04) + (Business Management × .09) + (Office Practices × −.03)]. Economic = [67.99 + (R × .12) + (I × .00) + (A × .15) + (S × .15) + (E × .11) + (C × .26) + (Ag × −.03) + (Nat × −.05) + (Adv × −.04) + (Mil × −.04) + (Mech × −.12) + (Sci × .03) + (Math × −.06) + (MSci × −.03) + (MServ × .05) + (Mus × −.12) + (Art × −.14) + (Writ × −.16) + (Teach × −.09) + (SServ × −.23) + (Athl × .02) + (Rel × −.16) + (PSpk × −.17) + (Law × −.03) + (Mrch × .08) + (Sales × −.08) + (Bus × .16) + (Off × −.11)]. Political = [45.61 + (R × −.01) + (I × −.08) + (A × .04) + (S × .04) + (E × −.02) + (C × .15) + (Ag × .00) + (Nat × .01) + (Adv × .02) + (Mil × .03) + (Mech × −.04) + (Sci × .00) + (Math × −.09) + (MSci × −.01) + (MServ × .01) + (Mus × −.15) + (Art × .01) + (Writ × −.06) + (Teach × −.01) + (SSer v × −.13) + (Athl × .15) + (Rel × −.19) + (PSpk × .19) + (Law × .20) + (Mrch × −.17) + (Sales × .13) + (Bus × −.01) + (Off × −.09)]. Aesthetic = [41.47 + (R × .04) + (I × .29) + (A × .03) + (S × −.11) + (E × .00) + (C × −.14) + (Ag × .02) + (Nat × .04) + (Adv × −.01) + (Mil × .00) + (Mech × .00) + (Sci × −.35) + (Math × .02) + (MSci × −.09) + (MServ × .08) + (Mus × .23) + (Art × .22) + (Writ × .08) + (Teach × .05) + (SServ × . 02) + (Athl × −.13) + (Rel × −.26) + (PSpk × −.14) + (Law × .00) + (Mrch × .08) + (Sales × .05) + (Bus × −.10) + (Off × .09)]. Social = [29.71 + (R × −.11) + (I × −.20) + (A × −.01) + (S × − .02) + (E × −.10) + (C × −.02) + (Ag × −.03) + (Nat × . 09) + (Adv × −.03) + (Mil × −.05) + (Mech × .12) + (Sci × −.01) + (Math × .05) + (MSci × .06) + (MServ × −.01) + (Mus × .08) + (Art × −.13) + (Writ × .03) + (Teach × .06) + (SServ × .44) + (Athl × .02) + (Rel × −.04) + (PSpk × .09) + (Law × −.06) + (Mrch × .14) + (Sales × .04) + (Bus × −.23) + (Off × .02)]. Religious = [11.31 + (R × .09) + (I × −.14) + (A × .02) + (S × −.10) + (E × −.03) + (C × −.21) + (Ag × .09) + (Nat × −.07) + (Adv × .05) + (Mil × −.02) + (Mech × −.08) + (Sci × − .02) + (Math × .08) + (MSci × .10) + (MServ × −.14) + (Mus × .03) + (Art × −.06) + (Writ × .00) + (Teach × −.01) + (SServ × .01) + (Athl × .01) + (Relig × .88) + (PSpk × .00) + (Law × −.04) + (Mrch × −.01) + (Sales × −.10) + (Bus × .10) + (Off × .12)].

Table 2
Correlations of the Study of Values (SOV) Scales and RIASEC Themes With External Criteria

Variable	SOV scale						RIASEC theme						*R*		
	Theoretical	Economic	Political	Aesthetic	Social	Religious	Realistic	Investigative	Artistic	Social	Economic	Conventional	SOV	Strong	Strong and SOV
Adjective Check List															
Dominance	.05	.11	**.23**	−.09	−.13	−.09	.03	.10	−.04	−.01	.14	.04	.23	.41	.42
Endurance	−.02	.09	.06	−.18	−.03	.08	−.02	.09	−.08	.09	−.02	.18	.20	.48	.51
Order	.04	.15	.08	**−.21**	−.11	.06	−.04	.07	−.17	.01	−.05	.18	.24	.50	.52
Intraception	−.08	−.08	−.16	−.08	**.24**	.10	−.09	.09	.08	**.21**	−.07	.09	.28	.44	.47
Nurturance	**−.25**	−.11	−.15	−.05	**.36**	.14	−.06	.04	.13	**.36**	.08	.12	.40	.52	.56
Affiliation	**−.22**	−.06	−.05	−.06	**.25**	.10	−.05	.03	.09	**.29**	.10	.09	.31	.48	.50
Autonomy	.15	.01	.16	.13	**−.21**	−.16	−.04	.03	.01	**−.23**	−.04	**−.20**	.33	.50	.51
Aggression	.11	.05	**.22**	.07	**−.26**	−.13	.03	−.04	−.04	**−.22**	.06	−.11	.34	.46	.49
Change	.03	−.10	−.06	.19	.09	−.14	.04	.05	.19	.01	.02	−.14	.25	.48	.50
Succorance	−.07	−.10	−.16	.07	.11	.09	−.03	−.05	.07	.07	−.03	.01	.17	.32	.33
Deference	−.15	.00	−.16	−.14	**.23**	.15	−.02	−.01	−.03	**.24**	.00	.19	.35	.54	.56
Self-Control	−.03	.02	−.13	−.13	.08	.13	−.02	.06	−.05	.11	−.08	.16	.23	.50	.52
Self-Confidence	.01	.06	.15	−.12	−.04	−.02	.04	.10	−.01	.06	.07	.01	.17	.39	.39
Creative Personality	.16	−.06	.02	.08	−.03	−.13	−.03	.07	.09	−.11	−.12	−.19	.24	.40	.44
Critical Parent	.15	.09	.18	.00	**−.26**	−.09	−.03	.01	−.09	**−.23**	.01	−.06	.30	.46	.47
Nurturing Parent	−.16	.00	−.02	−.14	**.20**	.09	−.02	.06	.01	**.28**	.06	.16	.29	.49	.50
High Origence/Low Intellectance	−.04	.04	.12	.11	−.05	−.12	.03	−.11	.06	.01	.12	−.06	.21	.46	.47
High Origence/High Intellectance	.04	−.18	−.15	**.30**	.00	−.05	−.10	−.10	.16	**−.24**	−.15	**−.33**	.34	.59	.61
Low Origence/Low Intellectance	−.16	−.01	−.10	−.08	**.27**	.05	−.04	−.02	.02	**.26**	.07	.13	.33	.46	.48
Low Origence/High Intellectance	.09	.14	−.02	**−.21**	−.05	.06	.01	.07	**−.20**	−.09	−.11	.12	.25	.44	.45
Multidimensional Personality Questionnaire															
Well-Being	−.08	−.12	−.01	−.04	.16	.07	.14	**.22**	.17	**.29**	**.20**	.18	.19	.42	.42
Social Potency	−.09	−.02	**.27**	.00	−.11	−.01	.02	.12	.18	.18	**.35**	.11	.32	.55	.57
Achievement	.05	.02	.02	−.15	−.01	.06	.15	**.35**	.09	**.20**	.14	**.25**	.16	.46	.49
Social Closeness	**−.27**	−.15	.03	−.02	**.29**	.09	.00	.04	.18	**.31**	**.21**	.12	.36	.47	.50
Stress Reaction	−.03	−.03	−.06	.02	.04	.04	−.01	−.10	.03	.00	.01	−.01	.07	.30	.30
Alienation	.14	.13	.10	−.10	**−.20**	−.03	.07	−.10	−.16	**−.20**	−.03	−.04	.22	.37	.39
Aggression	.15	**.24**	**.32**	−.12	**−.34**	−.14	.16	−.10	**−.24**	**−.31**	.03	−.08	.40	.54	.56
Control versus Impulsivity	.01	.08	−.02	−.13	.00	.05	−.06	.11	−.04	.11	−.02	**.20**	.16	.51	.52
Harm Avoidance	−.15	.05	−.06	−.05	.15	.04	**−.27**	−.16	−.05	.12	−.01	.13	.24	.64	.64
Traditionalism	−.19	−.03	−.08	**−.31**	.05	**.42**	.13	.08	−.04	**.24**	.05	**.23**	.49	.57	.62
Absorption	−.07	**−.28**	−.15	**.27**	.14	.04	.10	.17	**.40**	.19	.11	.01	.34	.50	.51
Family Environment Scale															
Achievement Orientation	.09	**.21**	.19	−.19	−.15	−.08	.06	.10	−.10	.01	.13	.17	.26	.35	.37
Intellectual–Cultural Orientation	−.08	−.17	−.12	.10	.12	.09	−.08	**.20**	**.25**	.17	.05	.05	.18	.37	.39
Moral–Religious Emphasis	**−.27**	−.06	−.08	**−.32**	−.04	**.59**	.09	−.02	−.04	**.22**	.12	.11	.64	.61	.69
Background data															
Books read in last year															
Science	**.21**	.04	.02	.04	−.06	−.18	.04	.17	.02	−.11	−.06	.00	.24	.35	.36

Table 2
(continued)

Variable	SOV scale						RIASEC theme						*R*		
	Theoretical	Economic	Political	Aesthetic	Social	Religious	Realistic	Investigative	Artistic	Social	Economic	Conventional	SOV	Strong	Strong and SOV
Plays/poetry/essays	−.14	−.15	−.16	.12	**.22**	.05	−.11	.02	**.24**	.18	.06	.01	.25	.42	.44
Love stories	**−.21**	−.09	−.15	.15	.16	.08	−.10	−.07	.17	.13	.04	.01	.26	.36	.37
Religious	−.15	−.07	−.08	−.04	.03	**.22**	.02	.04	.06	.13	.02	−.03	.23	.35	.36
How important is/are:															
Lots of money	.11	**.34**	**.26**	−.06	**−.25**	**−.26**	−.05	−.02	−.11	**−.20**	.07	.05	.40	.47	.51
Strong friendships	**−.21**	−.09	−.05	−.01	.19	.12	−.02	.03	.15	.17	.06	−.02	.24	.30	.36
Being a community leader	−.13	−.06	**.20**	−.16	.07	.08	.06	.15	.07	**.24**	**.25**	.19	.32	.45	.48
Future occupational importance of:															
Reading	**−.24**	**−.25**	−.06	.17	.15	.15	−.07	.01	**.29**	**.25**	.13	.05	.31	.41	.44
Writing	**−.21**	**−.24**	−.11	.19	.19	.10	−.13	−.02	**.31**	**.23**	.12	.00	.29	.45	.46
Social studies	**−.20**	**−.25**	−.08	.11	.16	.17	−.04	−.01	**.21**	**.21**	.11	.00	.28	.40	.44
Foreign language	−.18	−.15	−.07	.13	.13	.09	−.02	.05	**.21**	.17	.11	.06	.22	.28	.32
Computer science	.16	.17	.08	−.10	−.18	−.08	**.23**	.15	−.07	−.11	.02	.17	.21	.40	.41
Math	**.26**	**.25**	.02	**−.22**	**−.22**	−.05	**.20**	**.29**	−.13	−.08	.04	**.22**	.35	.52	.54
Physics	**.27**	.12	.06	−.16	−.14	−.08	**.28**	**.38**	−.10	−.09	−.02	.11	.28	.56	.57
Chemistry	.16	.08	.05	−.15	−.04	−.06	.16	**.40**	−.06	.06	.08	.12	.20	.54	.56
Biology	.07	−.05	.00	−.07	.06	−.01	.09	**.39**	.06	.15	.05	.08	.16	.58	.58
Tinker as a young child?	.19	.14	.10	−.09	**−.24**	−.05	.18	.10	−.18	**−.20**	−.08	−.03	.26	.42	.45
Tinker now?	**.31**	**.25**	.09	−.18	**−.29**	−.11	**.37**	.17	**−.27**	**−.24**	−.07	.08	.37	.61	.62
Preference for science courses	**.34**	**.25**	.06	**−.34**	−.14	−.10	.16	**.25**	**−.33**	−.16	−.08	.16	.44	.58	.60
Preference for humanities courses	**−.31**	**−.29**	−.08	**.30**	.18	.13	−.19	**−.20**	**.32**	.18	.06	−.11	.41	.54	.55
Ability tests															
SAT Verbal	.02	−.12	−.19	**.21**	.00	.03	−.11	.19	**.29**	.10	.06	.02	.27	.57	.58
SAT Math	.15	.04	−.06	−.05	−.10	.01	−.07	**.27**	.02	−.02	.03	.17	.18	.57	.58
Mechanical Comp.	**.26**	.15	−.02	−.08	**−.23**	−.05	**.23**	.08	**−.20**	**−.28**	−.11	−.06	.31	.52	.53
Mental Rotations	**.25**	.11	.04	−.09	−.18	−.09	**.20**	.11	−.17	**−.21**	−.07	.00	.27	.42	.44
Raven's Advanced Matrices	.06	−.06	−.07	.06	−.01	.00	.09	.14	.06	.00	−.04	.08	.13	.33	.35

Note. For external criterion variables, *n*s ranged from 670 to 695 (except $n_{SAT\ Math}$ = 331, and $N_{SAT\ Verbal}$ = 310). Correlations with external criteria of .20 or higher are boldfaced. For *N*s greater than 670, correlations of .10 or higher are statistically significant at $p < .01$. Strong = Strong Interest Inventory; SAT = Scholastic Aptitude Test.

Extrinsic Convergent Validation

Although observing impressive cross-validation coefficients is an important first step for establishing the interchangeability of an assessment procedure (Wiggins, 1988), high correlations between measures do not necessitate comparable external correlational profiles (Lubinski & Dawis, 1992; McCornack, 1956). To ascertain whether actual and forecasted SOVs are conceptually equivalent and empirically interchangeable, coupling traditional cross-validation applications with extrinsic convergent validation analysis (Fiske, 1971) provides more compelling conclusions than does the former alone. Following Fiske (1971), two measures should not be considered conceptually equivalent or empirically interchangeable until they display corresponding patterns in their correlational profiles. Ideally, this should be documented over a wide range of criteria (Lubinski & Dawis, 1992, p. 22), spanning as much of the convergent–discriminant range as possible with distinct data

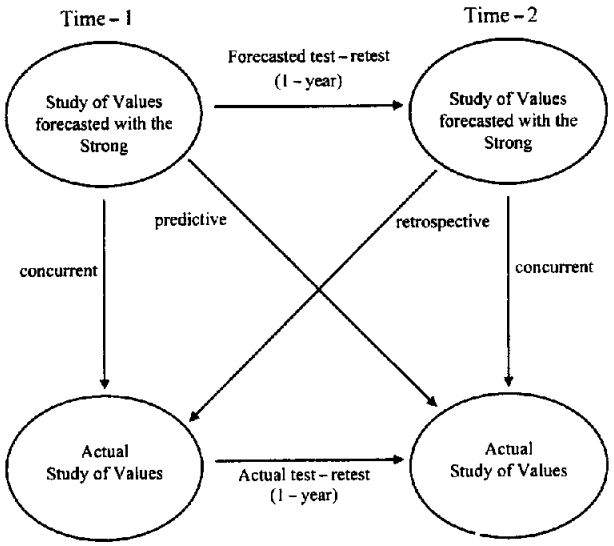

Figure 1. Cross-validation patterns.

sources. The previous cross-validation analyses revealed that Strong-based regression equations reliably reproduced inter- and intraindividual differences provided by SOV assessments. To speak more definitively to the validity of these Strong-based regression equations, however, we needed to examine the extent to which such forecasts covaried with external criteria in ways paralleling the actual SOV. Finding similar correlational profiles across actual and forecasted SOVs would provide ideal evidence that our regression equations produced valid, empirically interchangeable estimates of actual SOV scores.

Each person in the calibration sample had a total of four SOVs (actual and forecasted SOVs at Times 1 and 2). The 59 external criterion variables were correlated with each of the four SOV sources, for all six SOV scales. The resulting correlations were examined and, across corresponding themes, the profiles were quite similar over all four SOV sources.[9] To quantify this observation, we correlated pairs of correlational profiles. Averaged over the six SOV themes (following Figure 1), the pairs of profiles correlated as follows: actual test–retest = .89, forecasted test–retest = .94, concurrent cross-validation at Time 1 = .89, concurrent cross-validation at Time 2 = .93, predictive cross-validation = .89, and retrospective cross-validation = .88. Once again, there was a remarkable degree of consistency across all profile pairs, with the consistency of correlations between the two forecasted SOVs being the highest.

Discussion

Study 1 provided evidence for the construct validity of the Strong and the SOV for intellectually talented young adolescents. Convergent and discriminant correlational patternings of both instruments, across a wide array of external criteria, conformed to what one might expect to see in older populations. Evidence also suggested that for the external criteria we examined, the Strong provided incremental validity over the SOV, however, the reverse was not true. In fact, the Strong was so encompassing that perhaps a closer look at the traditional distinctions between interest and value assessment is warranted (at least with respect to the Strong and the SOV). We offer the analytic procedure we used to arrive at this conclusion as a general methodological approach for examining Dawis's (1992) concern over the amount of redundancy in psychological assessment tools.

Counseling and Theoretical Implications

Our results indicate that conventional preference questionnaires, initially designed for older students, manifest the same kinds of external relationships for young gifted adolescents that one would anticipate with older individuals. Given the amount of construct validity revealed here, practitioners might find these assessment tools valuable for helping gifted youth in educationally and vocationally relevant problem solving as well as for imparting better self-understanding. In particular, the results from the current study suggest that person–environment models, such as TWA, may be effectively used in refining the educational and vocational counseling offered to intellectually talented adolescents. This is especially relevant because these individuals begin thinking about career planning earlier than their age-mates do (Kerr & Erb, 1991; Milne, 1979; Silverman, 1993; Willings, 1986).

Theoretically, the ability/preference constellations manifested in Study 1 are of particular interest, because both differential (Humphreys, Lubinski, & Yao, 1993) and experimental (Kimble, 1984) psychologists have presented data indicating a psychological reality to C. P. Snow's (1967) "two cultures" (i.e., "humanists" and "scientists"; Lubinski, 1996). In Study 1, a patterning indicative of preferences for and nonlinguistic ideation about things versus preferences for and linguistic ideation about people emerged. Such ability–preference clusterings also have been noted in recent reviews involving more normative adult samples (Ackerman, 1996; Ackerman & Heggestad, 1997).

If optimal development unfolds, in part, by taking into account the enduring and most salient features of one's individuality (Scarr, 1992, 1996; Scarr & McCartney, 1983), the present findings may be helpful in facilitating positive development in gifted youth. For example, profiling abilities among the gifted to provide educational opportunities according to competence (as opposed to chronological age) has enjoyed much success (Benbow & Lubinski, 1996; Benbow & Stanley, 1996). Perhaps preference assessments could aid practitioners in refining how they approach gifted students with information about their abilities. For example, they could talk to gifted students about their exceptional capabilities (the intensity of their satisfactoriness) as well as about what they most enjoy (the importance of satisfaction). This could ultimately lead to enhancement of students' personal development more generally by helping students achieve a

[9] To conserve space, we do not present extrinsic validation profiles for Studies 1 and 2. These data are available from the authors.

Table 3
Interindividual Test–Retest and Cross-Validation Analyses

Study of Values theme	One-Year Study of Values test–retest		Cross-validation analyses							
			Concurrent at Time 1		Concurrent at Time 2		Predictive (Time 1 to Time 2)		Retrospective (Time 2 to Time 1)	
	Actual	Forecasted	28[a]	6[b]	28[a]	6[b]	28[a]	6[b]	28[a]	6[b]
Theoretical	.75	.77	.77	*.75*	.74	*.75*	.70	*.70*	.70	*.70*
Economic	.66	.76	.70	*.66*	.64	*.66*	.60	*.61*	.53	*.53*
Political	.68	.80	.73	*.70*	.75	*.71*	.68	*.65*	.64	*.59*
Aesthetic	.68	.80	.77	*.74*	.69	*.68*	.67	*.65*	.63	*.63*
Social	.63	.77	.58	*.57*	.59	*.60*	.55	*.54*	.58	*.60*
Religious	.79	.74	.78	*.78*	.78	*.77*	.72	*.73*	.66	*.64*
M	.70	.77	.72	*.70*	.70	*.70*	.65	*.65*	.62	*.62*

Note. Figure 1 diagrams these analyses. RIASEC = realistic, investigative, artistic, social, economic, conventional; Strong = Strong Interest Inventory.
[a]Using all 28 Strong scales (RIASEC and 22 Basic Interests). [b]Using the best six Strong scales (in italics).

more comprehensive understanding of themselves, others, and how they conceptualize their reactions to contrasting learning and work environments. By successively profiling enduring behavioral characteristics through normative assessments, counselors may position themselves to make some compelling and specific reflections and suggestions based on each student's unique idiographic constellation of personal attributes.

This should come as good news to counselors and educators trying to manage the huge literature on multipotentiality (see Achter, Benbow, & Lubinski, 1997; Achter et al., 1996) as well as indecisiveness among certain subsets of intellectually talented students. The educational and vocational paths genuinely amenable to students, traditionally classified as "multipotential" from an ability standpoint, are likely to be refined by attending to their preferences. By considering abilities and preferences conjointly, counselors are better positioned to identify more optimal sets of environments more individually tailored to an individual's particular proclivities and personal point of view.

Finally, our analyses converged on the conclusion that interest and values assessments may be reduced to assessing the former. If this is true, the three personal-attribute domains examined by Achter et al. (1996), which illustrate the amount of diversity within gifted populations (i.e., abilities, interests, and values), could be reduced to two assessment protocols without the loss of any relevant psychological information. Yet, a caveat is in order, because we can envision situations in practice in which an ipsative assessment is the preferred modality. For example, it is possible that practitioners, working one-on-one with certain clients, may be able to cogently argue that a forced-choice ipsative formatting is particularly useful for in-session self-exploration. One example could be the clarification of competing valences at comparable intensities (e.g., approach–approach or avoidance–avoidance conflicts). Examples such as this, however, are not sufficient to justify depicting the SOV as a domain-specific applied psychological research tool capable of complementing the Strong through the assessment of distinct psychological nuances.

Table 4
Means, Standard Deviations, and Quartiles (Q) of Intraindividual Test–Retest and Cross-Validation Correlations Estimated Parametrically and Nonparametrically

Statistic	One-Year Study of Values test–retest		Cross-validation analyses			
	Actual	Forecasted	Concurrent at Time 1	Concurrent at Time 2	Predictive (Time 1 to Time 2)	Retrospective (Time 2 to Time 1)
r						
M	**.69**	**.77**	**.73**	**.72**	**.69**	**.60**
SD	.29	.28	.27	.30	.31	.37
Q_1	.55	.68	.64	.62	.56	.41
Q_2	**.76**	**.87**	**.85**	**.83**	**.81**	**.70**
Q_3	.91	.95	.92	.92	.91	.88
ρ						
M	**.62**	**.72**	**.66**	**.66**	**.65**	**.54**
SD	.32	.30	.29	.32	.35	.38
Q_1	.42	.57	.51	.49	.55	.31
Q_2	**.69**	**.79**	**.74**	**.76**	**.77**	**.60**
Q_3	.85	.92	.88	.89	.89	.83

Note. Figure 1 diagrams these analyses. Mean and median correlations are boldfaced.

Measuring Constructs Through Ipsative Versus Normative Scaling

Two other issues emanating from the results of Study 1 involve ipsative scaling and the psychological nature of interest and value constructs themselves. With respect to the latter, if interests and values do indeed have psychologically distinct aspects, their distinctiveness should manifest itself in the context of external criteria; otherwise, there is no point in speaking about two classes of concepts when parsimony suggests that one will do. Clearly, the Strong met this requirement for the criteria we used, but the SOV fell somewhat short. Could all the psychological import found in measures of educational/vocational values (like the SOV) be subsumed by the Strong? Although this conclusion is premature, future research should direct itself toward this possibility (Rounds, 1990).

Perhaps the conspicuous differences between ipsative and normative assessment procedures nevertheless result in psychologically trivial differences with respect to the psychological dimensions assessed by these procedures. Indeed, such conspicuous differences in scaling may have attenuated the perceived need to ask this question. Yet, if interests and values are to be distinguished, scales purporting to assess each should capture their unique aspects, empirically, by manifesting incremental validity in the prediction of relevant criteria. This is paramount for documenting their applied psychological significance. It appears that much of the ipsativity offered by the SOV, and perhaps all of its psychological significance, may be gleaned normatively through Strong-based regression equations. Of course, these results are not definitive, inasmuch as all relevant criteria were not included in this study. Yet, we believe they are comprehensive enough to warrant consideration of the possibility. To further support this idea, consider the following. Whereas the sum of the six actual SOV themes was always 240, the sum of the six forecasted SOV themes ranged from 239.6 to 239.8. It seems, therefore, that the Strong is able to functionally reproduce the ipsativity of the SOV (i.e., as scores increase on one forecasted SOV theme, scores decrease on other forecasted SOV themes).

Again, just to be clear: Though the Strong appears to absorb the psychological import assessed by the SOV, this does not suggest that individual differences assessed by the SOV are unimportant—quite to the contrary (see Tables 1 and 2). It does suggest, however, that the six SOV dimensions (and much of their ipsativity) may be estimated through the Strong. The SOV has always focused on intraindividual differences, whereas the Strong has concentrated on interindividual differences, and yet both clearly assess common constructs. Whether ipsative assessments add incremental validity to normative scales in other psychological contexts is a question for further research.

A General Methodology

As mentioned earlier, Dawis (1992) expressed concern over the degree of redundancy that exists in instruments currently being used by counseling psychologists. Our findings in Study 1 exemplify this problem. We examined the problem of instrumentation (the Strong and the SOV) overlap by combining extrinsic convergent validation with conventional cross-validation techniques. We found that not only is Dawis's concern legitimate, but it may be operating in some of the more well-known instruments, not just the newly constructed instruments that Dawis highlighted. Given that this problem pervades many psychological domains, we offer the methodology we used here as one possible way to examine the redundancy in psychological measures. Ideally, it could result in a more parsimonious (less cluttered) subset of generic reference dimensions in counseling practice as well as in psychological science more generally. Our approach has an advantage over conventional multivariate procedures aimed at analyzing common variance. Analyzing the magnitude of between-instruments redundancy through conventional common variance techniques (e.g., canonical correlation or factor analysis) is less optimal than the procedures used here because scale uniqueness (i.e., scale specificity) could contribute incremental validity to relevant external criteria. Common variance procedures, by definition, jettison each scale's unique variance (specific variance + error variance) and focus on characterizing the dimensionality of common variance (or what variables share). This is why combining extrinsic convergent validation with conventional cross-validation techniques is an especially apt procedure for capturing unique aspects of psychological import that each assessment procedure holds. For innovative assessment procedures to make a contribution to psychological science, they must provide researchers with something they do not already have.

Study 2

Study 1 provided support for the construct validity of the Strong and the SOV for intellectually talented young adolescents. Cross-validation and extrinsic convergent validation analyses suggested that SOV assessments may be generated with precision and fidelity through Strong-based regression equations. Those results were, however, based on only one sample (albeit over two time-frames). Study 2 was a generalization probe in which we assessed the robustness of the aforementioned regression equations by applying them to a large sample of graduate students enrolled in top engineering, mathematics, and natural and physical science departments across the United States. This adult sample served as an independent calibration sample on which the regression equations based on gifted adolescents from Study 1 could be evaluated. Given that a prominent goal of our longitudinal research is to identify early individuals with the potential to excel in math and science disciplines (Lubinski & Benbow, 1994), top math and science graduate students are, in many ways, ideally suited for conducting a multiple-sample cross-validation design (Mosier, 1951; Murphy, 1984), or a generalizability probe, of this kind. Furthermore, if regression equations developed on gifted 13-year-olds manifest an impressive degree of cross-validation in exceptionally able, high-achieving adults, we would be in an excellent position to suggest that teaming ability and prefer-

ence assessments in counseling gifted adolescents is likely to result in applied psychological benefits. In short, in Study 2 we examined whether the adolescent covariance structure uncovered in Study 1 is sufficiently mature such that it may validly generate adult forecasts.

Method

Participants

The participants for this study were identified by SMPY in 1992 (Lubinski & Benbow, 1994; Lubinski, Benbow, Eftekhari-Sanjani, & Halvorson, 1998). They were students in the top 20 U.S. graduate training programs in engineering, mathematics, and natural or physical sciences (according to Gourman, 1989, and the National Research Council).[10] The majors pursued by these students included biochemistry, cellular and molecular biology, chemistry, computer science, engineering, mathematics, and physics. Participants who completed both the Strong and the SOV were included. This sample consisted of 695 participants (358 men and 337 women; 81.3% Caucasian, 9.5% Asian American, 2.4% African American, 3.2% Hispanic, .1% Native American, 2.9% other, and .6% not responding) between the ages of 23 and 30. Finally, because there tended to be a high male/female gender disparity in these departments, we implemented the following procedure. Departments were asked to secure the names of all 1st- and 2nd-year graduate students who qualified for and wished to participate in this study. This invariably resulted in a larger number of men than women. Of these, we took all the women and randomly sampled the same number from the male pool. This resulted in a comparable ratio of men to women even though the population ratio is greater than 4:1.

Instruments

The instruments we used in Study 2 are described in *Instruments* in the report of Study 1: the Strong, SOV, ACL (20 scales), and FES (3 scales). The participants also completed a background questionnaire, developed by SMPY. Six variables similar to those in Study 1's questionnaire were used: "How important is lots of money in your life? Strong friendships? Being a community leader?"; "What age did you become interested in math/science as a career?"; and "To what extent were you involved with 'tinkering' with equipment, mechanical gadgets, or construction games as a young child? As an adolescent?" Finally, Graduate Record Examination (GRE) Verbal, Quantitative, and Analytical scores were used here.

Procedure

Department chairs and heads were contacted in the spring of 1991 and asked whether SMPY could survey students in their department. Students were required to be American-born citizens, and approximately half of these students were female. Students who agreed to participate were sent a packet of questionnaires in April 1992. They mailed the completed packet back to SMPY researchers in a self-addressed, stamped envelope and received $15 in return. Because all participants were volunteers, the response rate was more than 99%.

Results

Cross-Validation

A benchmark. To evaluate the performance of Study 1's regression equations across more than one benchmark, we performed an independent cross-validation analysis within our graduate student sample by dividing them into two random samples (stratified by gender and major).[11] These constituted screening (n = 545) and calibration (n = 150) samples. The screening sample was used to develop a set of graduate-based regression equations, and the calibration sample was used to cross-validate these newly formed equations. These two samples were used only to generate a baseline against which we could compare the performance of the regression equations we built with the gifted adolescents. They allowed us to examine how our Study 1 regression equations performed relative to regression equations built on a sample of graduate students selected from the same population.

Interindividual analyses. The cross-validation coefficients, when Study 1 regression equations were used, are presented in Table 5. Comparing the cross-validation coefficients in Table 5 with those in Table 3 shows that the coefficients are comparable, other than a slight difference for Theoretical (which was anticipated because of the ceiling effects for math and science graduate students on this scale). Table 5 also contains cross-validation coefficients for the regression equations built with the graduate student screening sample. Clearly, the cross-validation coefficients were comparable to those from the adolescent sample. Interestingly, the two coefficients associated with Theoretical were nearly equivalent as well. It appears that the equations developed on the gifted adolescents hold just as much forecasting efficiency as those developed on the graduate students themselves! That is, the Study 1 regression equations generalize to top math and science graduate students and may be used to reliably generate their SOV profile.

Intraindividual analyses. We examined the intraindividual interchangeability of the actual- and forecasted-SOV profiles by evaluating the extent to which rank ordering of themes in the two profiles covaried. As in Study 1, Pearson product–moment and Spearman rank-order correlations were computed, intraindividually, for all 695 graduate students, and the central tendency and dispersion of the resulting distributions of coefficients were examined. The mean and

[10] Schools that participated were California Institute of Technology, Cornell University, Harvard University, Johns Hopkins University, Massachusetts Institute of Technology, New York University, Northwestern University, Princeton University, Stanford University, State University of New York at Stony Brook, University of California at Berkeley, University of California at Los Angeles, University of California at San Diego, University of Illinois, University of Michigan, University of Pennsylvania, University of Washington, University of Wisconsin, and Yale University.

[11] Before we examined the cross-validation of our regression equations, we assessed the intercorrelational comparability of the SOV and the Strong for gifted adolescents and graduate students. We correlated the intercorrelations of the adolescents' 6 SOV and 28 Strong theoretical scales with those of the graduate students. The correlations between the two samples' corresponding entries were .92 and .85 for the SOV and the Strong, respectively. Overall, the internal structure of the SOV for the adolescents was comparable to that of graduate students, whereas the Strong's internal structure was slightly less isomorphic for the adolescents.

Table 5
Cross-Validation of Regression Equations Based on Gifted Adolescents and Graduate Students and Multiple Correlations of Strong Scales (RIASEC and 22 Basic Interests)

Theme	Cross-validation		*R* (based on graduate students' Strong scales)
	Gifted adolescents' equations (r_{cv})	Graduate students' equations (r_{cv})	
Theoretical	.58	.57	.64
Economic	.66	.69	.69
Political	.60	.60	.65
Aesthetic	.69	.63	.72
Social	.54	.57	.61
Religious	.75	.75	.78
M	.64	.64	.68

Note. Strong = Strong Interest Inventory; RIASEC = realistic, investigative, artistic, social, enterprising, conventional.

median rank-order correlations were .65 and .76, respectively (SD = .33, Q_1 = .50, Q_3 = .88), and mean and median product–moment correlations were .70 and .80, respectively (SD = .31, Q_1 = .60, Q_3 = .91). This implies that the actual- and forecasted-SOV profiles produce comparable rank orders of themes. These values are very similar to those observed in Study 1 (see Table 4) and provide further evidence that the Strong is able to reliably reproduce the SOV profile for graduate students in top math and science programs.

Extrinsic Convergent Validation

To evaluate the extent to which these equations produced valid estimates of the SOV, we conducted an extrinsic convergent validation analysis. Thus, we examined the extent to which actual and forecasted SOVs had the same correlational pattern over a set of 32 external criteria (20 ACL scales, 3 FES scales, 6 background variables, and 3 GRE scores). Examination of the correlational profiles (see Footnote 9) revealed that across the external criteria, corresponding actual and forecasted SOV themes produced similar profiles. To quantify the degree of similarity, we correlated these two profiles. Averaged over the six SOV themes, the pair of profiles correlated .93, the same as the highest correlation of actual versus forecasted SOV profiles observed in Study 1. These results suggest that as was the case in Study 1, actual and forecasted SOVs are interchangeable.

Discussion

Study 2 provided evidence that the Strong and the SOV assess similar individual differences among gifted adolescents and math and science graduate students. Regression equations built on gifted adolescents were cross-validated and, more important, found to relate to external criteria for these graduate students in ways similar to the actual SOV. These findings suggest that individual-differences information obtained with Strong and SOV assessments of gifted adolescents is similar to that obtained when these instruments are administered to adults. Although, admittedly, our generalization probe (cross-validation and external validation analyses) designed to support this idea involved a highly specific adult (math and science) sample, we nevertheless venture a more global generalization. Because of the heterogeneity of the gifted adolescents on whose data these regression equations were built and validated (both verbally and mathematically gifted adolescents were included), we hypothesize that these equations will generalize to and carry the same degree of validity for other groups of exceptional students engaged in conceptually demanding but less quantitatively dense disciplines (e.g., the humanities and law). Moreover, it would not surprise us if this generalization extends to other adult populations as well, and gains empirical support in future investigations.

General Discussion

For intellectually gifted young adolescents, across a wide array of external measures, the Strong and the SOV behave (both convergently and discriminantly) in accordance with prior expectations derived from decades of empirical research on older populations (Dawis, 1991; Harmon et al., 1994). Our findings on ability–preference clusterings mirror those recently reported by Ackerman (1996; Ackerman & Heggestad, 1997), who examined adult populations in his studies of academic and intellectual development. Interestingly, the ability–preference clusterings observed among adults and gifted young adolescents (verbal abilities covarying with preferences for the humanities and mathematical–spatial abilities covarying with preferences for scientific pursuits) actually have much in common with and provide verisimilitude for C. P. Snow's (1967) two cultures (humanist and scientists).

Other empirical support for the validity of assessing preferences in this special population recently accrued through a discriminant function analysis of educational outcomes (Achter, Lubinski, Benbow, & Sanjani, 1998). An independent sample of 432 SMPY participants, assessed with both the SAT and the SOV during the 1970s, at or before age 13, were followed up 10 years later and asked what area they had earned their 4-year degree in. These

majors were then categorized into three groups: "humanities," "math–science," and "other." Two discriminant functions significantly discriminated these three groups and, importantly, both the SAT and the SOV added incremental validity relative to each other. Function 1 was defined by high positive loadings on the Math SAT and SOV Theoretical theme and high negative loadings on the SOV Social and Religious themes (a "scientific orientation"). Function 2 was defined primarily by two high positive loadings on the Verbal SAT and the SOV Aesthetic theme (a "humanistic orientation"). Again, note the configural patterning highly indicative of C. P. Snow's ideas. Humphreys et al. (1993) have discussed how predicting group membership through discriminant function analyses (especially when conducted over long temporal gaps) can provide valuable information for educational and vocational counselors. This information is complementary to conventional validation designs, which aim to predict individual differences on relevant criteria, such as that presented here. It appears that the two types of approaches to validation—predicting individual differences and group membership—converge on the same positive conclusion: preference assessments hold applied utility for this special population.

Because of their precocity, intellectually gifted young adolescents appear to be developmentally ready to profit from assessment procedures initially designed for much older students and young adults. Furthermore, an examination of their profile variability in this study and others (Achter et al., 1996; Lubinski et al., 1995; Lubinski et al., 1996) reveals that the intellectually gifted are anything but a categorical type ("multipotential" or otherwise; see Achter et al., 1997). They manifest an enormous breadth of diversity (Winner, 1996): The range of individual differences within this special population (like all special populations) necessitates that counselors respond to these students as individuals, or "idiographically," in accordance with their strengths and relative weaknesses. One way of doing this is by taking into account the most salient features of their individuality. This underscores a prominent theme in counseling psychology: the need for multifaceted assessment and focusing on client strengths (Dawis, 1992) both between and within various domains of enduring psychological characteristics (Benbow & Lubinski, 1997). Unidimensional approaches are almost always sure to come up short by comparison. The present investigation reveals that some instruments that carry validity for this special population have appreciable overlap, such that practitioners may be in a position to reduce the number of assessment tools necessary to characterize key aspects of their psychological makeup.

Actually, the redundancy observed between interests and values in this investigation is but one example of what is likely to be seen with other measures in other settings (Dawis, 1992). We suggest a general methodological approach for uncovering the genuine magnitude of incremental validity offered by multiple assessment procedures. The method combines Fiske's (1971) idea of extrinsic convergent validation with conventional cross-validation techniques. Using this approach may result in a smaller subset of generic reference dimensions in counseling practice and research as well as throughout psychological science more generally. In the future, creators of innovative assessment procedures and novel measures should devote more attention to documenting that new assessment tools provide counselors with something important that they do not already have.

For better understanding of educational/vocational adjustment and discord, TWA is one model that has always stressed the importance of multifaceted assessments. At the very least, individual-differences information on variables relevant to both satisfactoriness (abilities) and satisfaction (preferences) is needed to help educational and psychological services provide optimal learning and work environments. Although counseling psychologists have known for some time that differential expectations in learning should be tailored to ability level and pattern, preferences have only recently been evaluated for their validity in the context of relevant criteria for gifted youth. For intellectually talented young adolescents, the longitudinal stability of their educational/vocational preferences has gained empirical support (Lubinski et al., 1995; Lubinski et al., 1996), and their incremental validity, relative to abilities, for predicting educational outcomes has just emerged (Achter et al., 1998).

Given recent advances documenting the range and stability of preferences among gifted youth and the present findings, it seems reasonable to suggest that conventional psychological assessments of this special population would profit from a broadening of their scope to include these important nonintellectual personal attributes. This might be accomplished with preexisting, well-known questionnaires (e.g., the Strong and the SOV) initially designed for somewhat older individuals. The intent is not to pigeonhole students, but rather to capture general trends in their developmental trajectory and then help them see that for themselves. In addition to enhancing applied services for gifted students, such as educational programming and vocational counseling, assessing their preferences might have the potential for refining educational research, particularly with respect to uncovering relevant student characteristics (e.g., affective and conative functions relevant to learning; R. E. Snow, Corno, & Jackson, 1996) for tailoring more ideal learning opportunities. Indeed, a more comprehensive psychological profiling may be useful for imparting important personal information to these students, to help them take a more active role in building learning environments for themselves (Tyler, 1961, 1992; Williamson, 1965). This idea may be conceptualized in developmental terms through Scarr's (1996) work on building niches. According to Scarr, optimal positive development unfolds when one structures learning opportunities, whether "actively" (by oneself) or more "passively" (through interventions), in ways congruent with one's unique constellation of abilities, personality, and preferences.

References

Achter, J. A., Benbow, C. P., & Lubinski, D. (1997). Rethinking multipotentiality among the gifted: A critical review and recommendations. *Gifted Child Quarterly, 41,* 5–15.

Achter, J. A., Lubinski, D., & Benbow, C. P. (1996). Multipotentiality among the intellectually gifted: It was never there and already it's vanishing. *Journal of Counseling Psychology, 43,* 65–76.

Achter, J. A., Lubinski, D., Benbow, C. P., & Sanjani, H. (1998). *Assessing vocational preferences among gifted adolescents adds incremental validity to abilities: A discriminant analysis of educational outcomes over a 10-year interval.* Manuscript submitted for publication.

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

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121,* 219–245.

Allport, G. W., Vernon, P., & Lindzey, G. (1970). *Study of values.* Boston: Houghton Mifflin.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability among the intellectually talented: Their characterization, consequences, 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.

Benbow, C. P. (1992). Academic achievement in mathematics and science between ages 13 and 23: Are there differences among students in the top one percent of mathematical ability? *Journal of Educational Psychology, 84,* 51–61.

Benbow, C. P., & Lubinski, D. (1996). *Intellectual talent: Psychometric and social issues.* Baltimore: Johns Hopkins University Press.

Benbow, C. P., & Lubinski, D. (1997). Intellectually talented children: How can we best meet their needs? In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 155–169). Boston: Allyn & Bacon.

Benbow, C. P., Lubinski, D., & Suchy, B. (1996). The impact of SMPY educational programs from the perspective of the participant. In C. P. Benbow & D. Lubinski (Eds.), *Intellectual talent: Psychometric and social issues* (pp. 266–300). Baltimore: Johns Hopkins University Press.

Benbow, C. P., & Stanley, J. C. (1983). An eight-year evaluation of SMPY: What was learned? In C. P. Benbow & J. C. Stanley (Eds.), *Academic precocity: Aspects of its development* (pp. 205–214). Baltimore: Johns Hopkins University Press.

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.

Bennett, G. K. (1969). *Bennett Mechanical Comprehension Test: Forms S and T (test manual).* New York: Psychological Corporation.

Bennett, G. K. (1994). *Bennett Mechanical Comprehension Test* (2nd ed.). San Antonio, TX: Psychological Corporation.

Cattell, R. B. (1965). *The scientific analysis of personality.* New York: Penguin Books.

Clemans, W. V. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs.* (Supplement, Whole No. 14).

Cohn, S. J. (1991). Talent searches. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 166–177). Boston: Allyn & Bacon.

Dawis, R. V. (1991). Vocational interests, values, and preferences. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 2, pp. 833–871). Palo Alto, CA: Consulting Psychologists Press.

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

Dawis, R. V., & Lofquist, L. H. (1984). *A psychological theory of work adjustment: An individual differences model and its applications.* Minneapolis: University of Minnesota Press.

Duffy, E., & Chrissy, W. J. E. (1940). Evaluative attitudes as related to vocational interests and academic achievement. *Journal of Abnormal and Social Psychology, 35,* 226–245.

Ferguson, L. W., Humphreys, L. G., & Strong, F. W. (1941). A factorial analysis of interests and values. *Journal of Educational Psychology, 32,* 197–204.

Fiske, D. T. (1971). *Measuring the concepts of personality.* Chicago: Aldine-Atherton.

Gough, H. G., & Heilbrun, A. B. (1983). *The Adjective Checklist manual.* Palo Alto, CA: Consulting Psychologists Press.

Gourman, J. (1989). *The Gourman report: A rating of graduate and professional programs in American and international universities.* Los Angeles: National Education Standards.

Hansen, J. C., & Campbell, D. P. (1985). *Manual for the SVIB-SCII.* Palo Alto, CA: Consulting Psychologists Press.

Harmon, L. W., Hansen, J. C., Borgen, F. H., & Hammer, A. L. (1994). *Applications and technical guide for the Strong Interest Inventory.* Palo Alto, CA: Consulting Psychologists Press.

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

Holland, J. L. (1985). *Making vocational choices.* Englewood Cliffs, NJ: Prentice Hall.

Huba, G. J., & Hamilton, D. L. (1976). On the generality of trait relationships: Some analyses based on Fiske's paper. *Psychological Bulletin, 83,* 868–876.

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.

Kerr, B. A., & Erb, C. (1991). Career counseling with academically talented students: Effects of a value-base intervention. *Journal of Counseling Psychology, 38,* 309–314.

Kimble, G. A. (1984). Psychologists' two cultures. *American Psychologist, 39,* 833–839.

Kuse, A. R. (1977). *Familial resemblances for cognitive abilities estimated from two test batteries in Hawaii.* Unpublished doctoral dissertation, University of Colorado, Boulder.

Laudeman, K. A., & Griffeth, P. (1978). Holland's theory of vocational choice and postulated value dimensions. *Educational and Psychological Measurement, 38,* 1165–1175.

Lofquist, L. H., & Dawis, R. V. (1991). *Essentials of person-environment-correspondence counseling.* Minneapolis: University of Minnesota Press.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Lubinski, D. (1996). Applied individual difference research and its quantitative methods. *Psychology, Public Policy, and Law, 2,* 187–203.

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. Subotonik & K. D. Arnold (Eds.), *Beyond Terman: Contemporary longitudinal studies of giftedness and talent* (pp. 255–281). Norwood, NJ: Ablex.

Lubinski, D., & Benbow, C. P. (1995). Optimal development of talent: Respond educationally to individual differences in personality. *Educational Forum, 59,* 381–392.

Lubinski, D., Benbow, C. P., Eftekhari-Sanjani, H., & Halvorson,

M. B. J. (1998). *Our future scientific leaders: How many can be identified by age 13?* Manuscript in preparation.

Lubinski, D., Benbow, C. P., & Ryan, J. (1995). Stability of vocational interests among the intellectually gifted from adolescence to adulthood: A 15-year longitudinal study. *Journal of Applied Psychology, 80,* 196–200.

Lubinski, D., & Dawis, R. V. (1992). Aptitudes, skills, and proficiencies. In M. Dunnette & L. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3, pp. 1–52). Palo Alto, CA: Consulting Psychologists Press.

Lubinski, D., Schmidt, D. B., & Benbow, C. P. (1996). A 20-year stability analysis of the Study of Values for intellectually gifted individuals from adolescence to adulthood. *Journal of Applied Psychology, 81,* 443–451.

McCornack, R. L. (1956). A criticism of studies comparing item-weighting methods. *Journal of Applied Psychology, 40,* 343–344.

Milne, B. G. (1979). Career education. In A. H. Passow (Ed.), *The gifted and talented: Their education and development* (pp. 246–254). Chicago: University of Chicago Press.

Moos, R. H., & Moos, B. S. (1986). *Family Environment Scale manual.* Palo Alto, CA: Consulting Psychologists Press.

Mosier, C. I. (1951). Problems with cross-validation. *Educational Measurement, 11,* 5–11.

Murphy, K. R. (1984). Cost–benefit considerations in choosing among cross-validation methods. *Personnel Psychology, 37,* 15–22.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Raven, J. C. (1973). *Advanced Progressive Matrices.* London: H. K. Lewis.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* London: H. K. Lewis.

Rounds, J. B. (1990). The comparative and combined utility of work value and interest data in career counseling with adults. *Journal of Vocational Psychology, 37,* 32–45.

Rounds, J. B., & Tracey, T. J. (1993). Prediger's dimensional representation of Holland's RIASEC circumplex. *Journal of Applied Psychology, 78,* 875–890.

Sanders, C. E., Lubinski, D., & Benbow, C. P. (1995). Does the Defining Issues Test measure psychological phenomena distinct from verbal ability?: An examination of Lykken's query. *Journal of Personality and Social Psychology, 69,* 498–504.

Sarbin, T. R., & Berdie, R. F. (1940). Relation of measured interests to the Allport-Vernon Study of Values. *Journal of Applied Psychology, 24,* 287–296.

Scarr, S. (1992). Developmental theories of the 1990s: Development of individual differences. *Child Development, 63,* 1–19.

Scarr, S. (1996). How people make their own environment: Implications for parents and policy makers. *Psychology, Public Policy, and Law, 2,* 204–228.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54,* 424–435.

Silverman, L. K. (Ed.). (1993). *Counseling the gifted and talented.* Denver, CO: Love.

Snow, C. P. (1967). *The two cultures.* Cambridge, England: Cambridge University Press.

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

Spranger, E. (1928). *Types of men.* New York: Stechert-Hafner.

Tellegen, A. (1982). *Brief manual of the Multidimensional Personality Questionnaire.* Unpublished manuscript, University of Minnesota, Minneapolis.

Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorders. *Psychological Inquiry, 4,* 122–130.

Tellegen, A., & Waller, N. G. (in press). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In S. R. Briggs & J. M. Cheek (Eds.), *Personality measures: Development and evaluation* (Vol. 1). Greenwich, CT: JAI Press.

Tracey, T. J., & Rounds, J. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin, 113,* 229–246.

Tyler, L. E. (1961). The future of vocational guidance. In M. S. Viteles, A. H. Brayfield, & L. E. Tyler (Eds.), *Vocational counseling* (pp. 59–70). Minneapolis: University of Minnesota Press.

Tyler, L. E. (1992). Counseling psychology—Why? *Professional Psychology: Research and Practice, 23,* 342–344.

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

Wiggins, J. (1988). *Personality and prediction.* Malabar, FL: Robert E. Kriger.

Williams, C. M. (1972). Occupational choice of male graduate students as related to values and personality: A test of Holland's theory. *Journal of Vocational Behavior, 2,* 39–46.

Williamson, E. G. (1965). *Vocational counseling: Some historical, philosophical, and theoretical perspectives.* New York: McGraw-Hill.

Willings, D. (1986). Enriched career search. *Roeper Review, 9,* 95–100.

Winner, E. (1996). *Gifted children.* New York: Basic Books.

Received September 15, 1997
Revision received May 27, 1998
Accepted June 1, 1998 ■