Journal of Vocational Behavior **61**, 139–184 (2002)
doi:10.1006/jvbe.2002.1878

MONOGRAPH

Convergent and Discriminant Validity of Five Interest Inventories[1]

Mark L. Savickas

Northeastern Ohio Universities College of Medicine

Brian J. Taber

Kent State University

and

Arnold R. Spokane

Lehigh University

This study investigated whether interest inventories that purport to measure the same constructs actually yield scores that correspond. The study examined the empirical relation of scores for similarly and same-named scales on five widely used interest inventories: the *Campbell Interest and Skills Survey*, the *Kuder Occupational Interest Survey–Form DD*, the *Self-Directed Search*, the *Strong Interest Inventory–Skills Confidence Edition*, and the *Revised Unisex Edition of the ACT Interest Inventory*. Comparisons were made among scores for (a) vocational interests measured by homogeneous, rationally based scales; (b) occupational interests measured by heterogeneous, criterion-based scales; and (c) self-efficacy for RIASEC tasks. The participants consisted of 80 women and 38 men employed as career counseling practitioners and professors. Results from analyses of multitrait–multimethod matrices indicated that similarly and same-named scales correlated moderately and that, with few exceptions, these matched scales demonstrated convergent and discriminant validity. These conclusions were interpreted by distinguishing between the linguistic explication and operational definition of constructs in theories of vocational and occupational interests. The implications of these interpretations were considered for both the science of vocational psychology and the practice of career counseling. Future research should investigate both the profile validity and the interpretive validity of interest inventories that yield scale scores derived from different scaling strategies.

---

[1] Howard E. A. Tinsley served as the action editor for this monograph.

Address correspondence and reprint requests to Mark L. Savickas, Department of Behavioral Sciences, Northeastern Ohio Universities College of Medicine, 4209 State Route 44, Rootstown, OH 44272-0095. E-mail: ms@neoucom.edu.

0001-8791/02 $35.00

Key Words: vocational interests; occupational interests; interest inventories; Self-Directed Search; Strong Interest Inventory; self-efficacy scales.

Interest inventories are among the most commonly used measures in career counseling practice and research (Spokane & Jacob, 1996). The most frequently used interest inventories display striking differences in scale type, item format, and norming procedures (Harmon, 1999). Despite these marked differences, some career counselors and researchers assume that the inventories should produce similar results, and they typically select a single inventory with which to measure the vocational interests of their clients or research participants. This assumption of homogeneity among similarly named scales should be at issue because of the differences in inventory construction. A rose is a rose by any name, but are artistic interests the same on any inventory? To date, we know little, if anything, about the similarity of scale scores among the current generation of interest inventories: the *Campbell Interest and Skills Survey* (CISS) (Campbell, Hyne, & Nilsen, 1992), the *Kuder Occupational Interest Survey–Form DD* (KOIS) (Kuder & Zytowski, 1991), the *Self-Directed Search* (SDS) (Holland, Fritzsche & Powell, 1994), the *Strong Interest Inventory* (SII) (Harmon, Hansen, Borgen, & Hammer, 1994), and the *Revised Unisex Edition of the American College Testing (ACT) Interest Inventory* (UNIACT-R) (ACT, 1995). Accordingly, the current study examined the construct validity of these inventories. Specifically, the study compared scores from similarly and same-named scales (hereafter referred to as matched scales) on these five inventories to determine the convergent and discriminant validity of the scales. The study also sought to reinvigorate a line of research that, although productive during the 1960s and 1970s, lay mostly dormant during the 1980s and 1990s, as shown in the following review of literature concerning the similarity of results obtained from the same individual responding to two or more interest inventories.

In the first study that we could identify on the similarity of interest inventory results, Triggs (1943) compared the *Kuder Preference Record* (KPR) and the *Strong Vocational Interest Blank* (SVIB) in a sample of 267 undergraduates. Triggs reported that the grouping scales (comparable to RIASEC theme scales) on each inventory correlated from a low of $r = .33$ on their respective Social Service scales to a high of $r = .69$ on their respective Enterprising scales (i.e., KPR Persuasive scale and SVIB Sales Manager scale). In a subsequent study that used a sample of 166 men from his previous study, Triggs (1944) concluded that SVIB occupational scales correlated positively with matched KPR scales, whereas negative correlations occurred between conceptually different scales. For instance, the SVIB Sales Manager scale correlated $r = .72$ to the KPR Persuasive scale and $r = -.38$ to the KPR Scientific scale. These two early studies supported the convergent and discriminant validity of the scales in that generation of the Kuder and Strong inventories.

Subsequent studies reported lower correlations between matched scales, with some studies noting higher correlations between conceptually dissimilar scales. For instance, King, Norrel, and Powers (1963) examined relations between matched

occupational scales on the KOIS and SVIB in a sample of 464 male undergraduates attending a college of business administration. King and his colleagues (1963) failed to replicate the results reported by Triggs (1943, 1944), finding instead lower average correlations of $r = .37$ for 14 same-named scales and $r = .39$ for 18 similarly named scales. They reported that some scales correlated higher to unrelated scales than to their matched scales. For example, the Physician scales on the SVIB and KOIS correlated $r = .41$, whereas the SVIB Physician scale and the KOIS Chemist scale correlated $r = .45$. Using a sample of 175 male undergraduates who presented voluntarily for career counseling, O'Shea and Harrington (1971) reported the same type of unexpected findings. The SVIB Engineering scale and several KOIS Engineering scales correlated $r = .18$ to $r = .45$, whereas the SVIB Engineering scale correlated higher to six other scales, including Auto Mechanic ($r = .52$) and Carpenter ($r = .47$). Overall, O'Shea and Harrington reported a median correlation of $r = .39$ for matched scales on the KOIS and SVIB. This moderate correlation resembled the average correlation of $r = .32$ between 27 matched scales of the KOIS and SVIB reported in a study by Wilson and Kaiser (1968). Johnson (1971) also compared scales on the KOIS and SVIB for 137 female undergraduates and reported similar results; median correlations between matched scales ranged from $r = .31$ to $r = .34$.

In the first of a series of programmatic studies comparing interest inventories, Zytowski (1968) investigated relations between matched scales on the SVIB, KOIS, and *Minnesota Vocational Interest Inventory* (MVII) (Clark, 1961) with data collected from 224 undergraduate males enrolled in military science courses. The median correlation for matched scales of the SVIB and KOIS was $r = .25$; for the KOIS and MVII, the median correlation was $r = .19$; and for the SVIB and MVII, the median correlation was $r = .08$. A subsequent study by Zytowski (1972a) correlated scores on 52 matched scales in the SVIB and KOIS for two separate samples. The first sample came from the study mentioned above, and the second sample consisted of 108 university counseling center clients. Using the traditional method, Zytowski computed median correlations of $r = .25$ in the first sample and $r = .39$ in the second. Zytowski then correlated the scale pairs for each participant in both samples. He used a frequency distribution to examine the scale correlations. This method of analysis revealed higher scale correlations between the two inventories: a median correlation of $r = .57$. Zytowski (1972b) further explored the comparability of the SVIB and KOIS by assessing their accuracy in classifying occupations. In Zytowski (1972b), 290 participants who were established in occupations had taken both inventories. Results of the comparison showed that the SVIB and KOIS correctly classified 65.5% and 71.1% of the participants, respectively. However, only 53% of the participants were correctly classified by both inventories. This study suggested that the two inventories worked equally well yet differently.

Two studies examined correlations between matched scales on the *Vocational Preference Inventory* (VPI) (Holland, 1977), the *California Occupational Preference System (COPS) Interest Inventory* (Knapp & Knapp, 1985), and the

Kuder General Interest Inventory–Form E (KGIS). A sample of 177 eighth-grade students participated in a study conducted by Best and Knapp-Lee (1982) that compared matched scales of the KGIS and COPS. The correlations ranged from $r = .21$ for the COPS Communication scale and the KGIS Literary scale to $r = .49$ for the KGIS Science scale and the COPS Science–Professional scale. The median correlation for the matched scales was $r = .42$. Similarly, Omizo and Michael (1983) compared matched scales on the VPI and COPS for 213 female undergraduate students. The comparison revealed modest correlations ranging from $r = .38$ for the VPI Conventional scale and the COPS Clerical scale to $r = .54$ for the VPI Realistic scale and the COPS Technical–Skilled scale. This same correlation was obtained between the VPI Investigative scale and the COPS Scientific–Professional scale. Approximately 50% of the sample had the same top three interests on both inventories.

One study examined how KOIS interpretations influenced relations between matched SVIB and KOIS scales. Carek (1972) asserted that an important factor in interpreting the KOIS is the rank order of the scales. He reasoned that by following the KOIS interpretive guidelines, the relations between matched scales on the SVIB would be higher. Accordingly, Carek correlated rank order of KOIS scores to SVIB standard scores in a sample of 120 males who requested career counseling services. He reported correlations that ranged from $r = .07$ between the Architect scales to $r = .76$ between the Minister scales, with a median of $r = .49$. The results of this analysis were then compared to the findings of O'Shea and Harrington (1971), Zytowski (1968), and Wilson and Kaiser (1968) to determine whether using this method would yield different results from those reported in previous studies. The hypothesis was not confirmed for all comparisons. A total of 27 correlations were significantly higher than those reported previously, 5 were significantly lower, and 19 comparisons did not yield statistically significant differences between correlations.

In sum, studies that have investigated the comparability of matched scales on two or more interest inventories have reported, at best, moderate correlations. In seven studies, the correlation coefficients for matched occupational scales were generally in the upper .30s, and in two studies, the correlations for homogeneous scales were $r = .49$ and $r = .54$. These results suggest that there seems to be only moderate agreement between interest inventories purporting to measure the same construct.

Why do matched scales of interest inventories have low to moderate correlations? Researchers have offered speculative answers to this question. Wilson and Kaiser (1968) thought that differences in norm groups may have influenced the results of their study. Different scoring procedures and item response inconsistency or unreliability of the inventories themselves were cited as possible explanations by Zytowski (1968, 1972). Johnson (1971) explored empirically possible reasons for such inconsistencies and reported that individuals with inconsistent profiles scored higher on the Imagination scale of the 16PF. He concluded that more imaginative participants were less concerned with providing consistent responses on both inventories and were perhaps careless in completing the inventories. Carek (1972)

accounted for the moderate correlations between matched scales by concluding that the KOIS and SVIB "measure different aspects of the same occupation" (p. 222).

Whether or not they measure different aspects of the same occupations, it seems certain that different inventories work equally well in predicting occupational entry. Kuder (1969) acknowledged that the KOIS and SVIB both accomplish the same goal yet emphasized the point that the occupational scales use different methods to meet this goal. It is these different methods, and not differences in accuracy of outcomes, that make low correlations inevitable.

Scale items on the SVIB differentiate members of an occupational group from a general reference group. This procedure for item selection removes from each scale the core of common interests that people share, leaving only the items unique to the occupational group. The resulting scale scores sum the item responses and are comparable from person to person. By contrast, scale scores on the KOIS are correlation coefficients that are comparable for one person in terms of rank order but not comparable from person to person. The KOIS lambda coefficient includes the respondent's similarity both to members of a specific occupational group and to people in general.

Kuder (1969) attributed the low correlations between matched occupational scales to the fact that SVIB scores exclude common interests, whereas KOIS scores include common interests. Strong (1943) claimed that the common core, which he partialled out of his occupational scales, was greater than the variance that remains in any scale, with each scale having a different proportion of common variance to specific variance. Kuder suggested that if the common variance removed from an SVIB scale is just 50%, then the maximum correlation possible between the matched KOIS and SVIB scales is reduced to the square root of 50, that is, $r = .71$. If the common variance is 60%, then the upper limit of correlation becomes $r = .63$. This upper limit is further reduced by the reliability of the scales, making low correlations inevitable yet still allowing the inventories to succeed equally well in predicting occupational entry. Of course, Kuder's argument applies only to heterogeneous, criterion-based occupational scales, and not the homogeneous vocational interest scales, leaving counselors to wonder about the comparability of results from the different types of scales on contemporary interest inventories.

Since the publication of the studies reviewed herein and Kuder's (1969) interpretation of their meaning, the interest inventories in question have undergone numerous revisions and new interest inventories have been published. Several interest inventories have adopted Holland's (1997) RIASEC typology to organize results. In addition to revisions and modifications of reporting methods, a new construct has been added to some interest inventories. Self-efficacy scales have been incorporated as an option on the Strong Interest Inventory and as an integral part of the Campbell Interest and Skills Survey. Given the development of the current generation of interest inventories, along with the hiatus in research on matched scales, we decided to investigate the issue of scale similarity on widely used inventories. The literature on the concurrent and predictive validity of these inventories

already shows their effectiveness in facilitating and predicting vocational choice and occupational entry.

The Current Study

Throughout the article, we use the phrase "vocational interests" to denote interests in school and work activities, and we use the phrase "occupational interests" to denote the similarity of an individual's vocational interests to the interests of a group of individuals employed in particular occupations. This study systematically compared five interest inventories at three levels of comparison: (a) vocational interests measured by homogeneous scales; (b) occupational interests measured by heterogeneous, criterion-based scales; and (c) self-efficacy for tasks associated with RIASEC themes. Specifically, the study assessed the convergent and discriminant validity of similarly and same-named scales on each inventory and considered the interchangeability of the inventories as measures of vocational and occupational interests. The approach taken in the current study directly evaluates the construct validity of the inventories as research instruments. It only indirectly addresses the issue of validity for counseling use (Cronbach, 1980) and whether the instruments are sufficiently interchangeable to be treated as parallel forms.

We approached the construct validation process by using Campbell and Fiske's (1959) data-analytic technique of the multitrait–multimethod (MTMM) matrix to compare the five methods of inventorying three sets of career constructs: vocational interests as indicated by RIASEC scales, occupational interests as indicated by occupational scales, and vocational self-efficacy as indicated by skills rating scales. In Campbell and Fiske's original description of the MTMM technique, the traits were measured by very different methods, for example, an objective inventory, a projective technique, and a structured interview. However, because we were studying the adequacy of inventories as a method for measuring vocational and occupational interests, we measured the three sets of career constructs using five different inventories. Of course, the actual inventorying methods used in the five measures are quite diverse. As described under Methods, the measurement methods used in these inventories differ in item content, item format, response scale, normative criteria, and scoring procedures.

METHODS

Participants

The participants consisted of attendees at a Society for Vocational Psychology conference on vocational interests. Attendees completed the five interest inventories (CISS, KOIS, SDS, SII, and UNIACT-R) before the conference and received their results at the conference. Of the 150 conference participants, 118 (80 women and 38 men) agreed to submit their interest inventory results for use in this study. The data they submitted were anonymous but did indicate their sex and occupation. Of these participants, 42.4% ($n = 50$) identified themselves as career counselors, 41.5% ($n = 49$) identified themselves as career counseling researchers and

professors, 10% ($n = 12$) listed their primary occupation as "other," and 5.9% ($n = 7$) did not specify their occupation.

Measures

The study compared matched scales on five of the most commonly used interest inventories: the Campbell Interest and Skills Survey (Campbell et al., 1992), the Kuder Occupational Interest Survey (Kuder & Zytowski, 1991), the Self-Directed Search (Holland et al., 1994), the Strong Interest Inventory–Form T317 Skills Confidence version (Harmon et al., 1994), and the Revised Unisex Edition of the ACT Interest Inventory (ACT, 1995). Complete technical data appear in the above-cited manuals and are briefly summarized herein.

Campbell Interest and Skills Survey. The CISS consists of 320 items that measure vocational interests in seven orientations (Influencing, Organizing, Helping, Creating, Analyzing, Producing, and Adventuring) and vocational self-efficacy beliefs that correspond to each interest scale. Scores are reported as T-scores. The seven Orientation Scales resemble the RIASEC typology (Holland, 1997) with the addition of a seventh scale, Adventuring. The Influencing orientation resembles the Enterprising type and encompasses occupations in which people lead and persuade others. Individuals who score high on this scale frequently want to take charge and enjoy being visible in the public. The Organizing orientation resembles the Conventional type and involves activities that require orderliness and planning. Individuals who score high on this scale enjoy detail and solving day-to-day problems associated with an organization. The Helping orientation resembles the Social type and involves personal service in which individuals express genuine concern for the well-being of other people. Individuals who score high on the Helping orientation typically enjoy close personal contact with others and may prefer occupations such as teaching and counseling. The Creating orientation resembles the Artistic type and involves activities such as acting, writing, musical performance, and other artistic endeavors. Individuals who score high in this orientation are likely to enjoy creating new products and ideas within artistic areas. The Analyzing orientation resembles the Investigative type and involves activities in which people work with data and numbers. Individuals who score high on this scale have a need to understand the world in a scientific sense. The Producing orientation resembles the Realistic type and involves practical activities that produce useful products. Individuals who score high on this scale generally enjoy manual work and like to see the results of their labor. The Adventuring orientation reflects a different aspect of the Realistic type, one that involves activities that require physical endurance, risk taking, and competition with others. Individuals who score high on this scale prefer physical activities and seek excitement.

The CISS also contains Basic Interest and Skill Scales, Occupational Scales, and Special Scales. The 29 Basic Interest and Skill Scales are organized within the 7 orientations and reflect the respondent's attraction to a specific area and sense of confidence about her or his ability to perform the activities in a specific area. The 29 scales report parallel interests and skills to provide more refined information

to the individual. The 58 Occupational Scales each measure resemblance to the interests and skills confidence of workers within a specific occupation. The CISS uses a 6-point Likert scale response format ranging from *strongly like* to *strongly dislike*.

Campbell et al. (1992) provided reliability information in the CISS manual. Internal consistency (alpha) coefficients for the Orientation Scales ranged from .86 to .93 for interests and from .76 to .89 for self-efficacy. Test–retest reliability for the Orientation Scales for a 3-month period ranged from $r = .81$ to $r = .88$ for interests and from $r = .75$ to $r = .84$ for skills. Internal consistency (alpha) coefficients for the Basic Interest and Skill Scales ranged from .69 to .92 for interests and from .62 to .87 for skills. Test–retest reliability within a 3-month period for the Occupational Scales ranged from $r = .73$ to $r = .92$ for interests and from $r = .73$ to $r = .84$ for self-efficacy.

Few validity studies of the CISS have been reported. Campbell and his colleagues (1992) established the content validity of the Orientation Scales through correlations of matched interest and skill scores. Correlations between the respective interest and skill Orientation Scales ranged from $r = .66$ to $r = .76$ with a median correlation of $r = .70$. Additional evidence for content validity of the Orientation Scales was established by plotting the interest means for each of 58 occupational samples along each of the Orientation Scales to show the accuracy of scale placement. The occupational samples associated with their respective Orientation Scales ranked higher within that scale than did occupational samples that had different orientations. These procedures were repeated to establish the content validity of the basic scales. Correlations between the skill and interest scores ranged from $r = .46$ to $r = .80$ with a median correlation of $r = .68$. The mean interest scores of 58 occupational samples were plotted in rank order from highest to lowest on their corresponding basic scales. Content validity for the occupational scales was established through correlations of the interests and skills of the Occupational Scales and the interests and skills on their respective Orientation Scales. As expected, the Occupational Scales correlated highest with their own Orientation Scales. Further validity evidence for the Occupational Scales has been demonstrated by determining the extent to which scales discriminated between those who were employed in a specific occupation and those who were not. The median validity index for the interest scales was 1.97 and for the skill scales was 1.82, meaning that those employed in a particular occupation had mean scores approximately 2 standard deviations above that general reference sample. Criterion validity evidence for the CISS appeared in a study by Hansen and Neuman (1999), who reported that the CISS had an overall hit rate of 72.6% for college students in relation to major selection.

Kuder Occupational Interest Survey–Form DD. The KOIS contains 100 sets of three activities. A respondent ranks the activities in each set from most preferred to least preferred. The profile contains 10 Vocational Interest Estimate Scales (VIEs), with scores reported as percentiles: Scientific, Artistic, Literary, Social Service, Musical, Outdoor, Computational, Clerical, Persuasive, and Mechanical. The manual explains how to convert VIEs into the RIASEC typology.

Preferences for Outdoor and Mechanical activities reflect the Realistic type. Scores on the Outdoor and Mechanical scales are summed and divided by 2 to obtain a score for the Realistic type. Outdoor preferences involve activities such as growing plants, raising animals, forestry, and farming. Mechanical preferences refer to working with tools and machines and activities such as making repairs. Literary, Artistic, and Musical preferences reflect the Artistic type. Scores on the Literary, Artistic, and Musical scales are summed and divided by 3 to obtain a score for the Artistic type. Literary preferences involve activities such as reading and writing. Artistic preferences involve creating and designing activities such as painting and sculpting. Musical preferences involve activities such as attending concerts and playing a musical instrument. Clerical and Computational preferences reflect the Conventional type. Scores on the Computational and Clerical scales are summed and divided by 2 to obtain a score for the Conventional type. Clerical preferences involve clearly defined tasks that require precision and accuracy. Computational preferences involve an interest in working with numbers. Scientific preferences reflect the Investigative type and involve understanding and solving problems in the physical world. Social Service preferences reflect the Social type and involve helping people. Persuasive preferences reflect the Enterprising type and involve convincing people about a particular point of view, promoting projects, and/or selling things.

The KOIS also contains 109 Occupational Scales and 40 College Major Scales, with scores reported as lambda coefficients that indicate how well an individual's responses resemble the criterion group for each occupation and college major. The *Kuder General Manual* reports that Occupational Scales and College Major Scales have test–retest reliabilities in the .90s. The VIEs have a profile stability of $r = .80$ for a 2-week interval (Diamond, 1990). The manual also documents the ability of the KOIS to discriminate between groups as evidence of concurrent validity. A study by Zytowski (1976) with 884 men and women examined the predictive validity of the KOIS. Zytowski reported that 12 to 19 years after the participants had taken the KOIS, 50% were working in occupations that coincided with their KOIS results. Using 206 persons whose occupations were predicted from one of the KOIS twin scales (occupational scales normed separately for men and women) from the 1976 sample, Zytowski and Laing (1978) reported that 43% were in occupations ranked in the top five of their gender norm and 50% were in occupations ranked in the top five of the other-gender normed scales.

Self-Directed Search–Form R. The SDS is a self-scored inventory that measures resemblance to RIASEC types. The inventory can be self-administered and self-interpreted with the use of the *Occupations Finder* (Holland, 1994). The assessment booklet contains four sections: Activities (6 scales with 11 items each), Competencies (6 scales with 11 items each), Occupations (6 scales with 14 items each), and Self-Estimates (2 sets of 6 ratings). Raw scores in the four sections are summed to indicate a person's resemblance to the RIASEC vocational personality types. The response format for the items is *like* or *dislike* for the Activities section; it is *yes* or *no* for the Competencies section and Occupations section; and for

Self-Estimates, respondents provide two ratings for each type of ability and skill using a 7-point numerical rating scale where 7 = *high*, 4 = *average*, and 1 = *low*. Spokane and Holland (1995) reported that for the 1994 revision, the summary scales reliability ranged from .90 to .94, internal consistency for each of the separate sections ranged from .72 to .92, and test–retest reliability ranged from $r = .76$ to $r = .89$ for a period of 4 to 12 weeks.

The *Technical Manual for the Self-Directed Search* (Holland et al., 1994) provides extensive information concerning validity of the scales on previous versions of the SDS. Although results of studies with previous versions of the SDS support the concurrent and predictive validity of previous versions, no such studies of the 1994 version have been reported. Correlations between the scales on the 1994 version have shown that, with few exceptions, the relations occur in the theoretically expected magnitude and direction.

Strong Interest Inventory Skills–Confidence Edition. The SII contains 317 items scored in three sets of scales, with results reported as T-scores. The first set of scales are the six General Occupational Themes (GOTs) based on the RIASEC types. The second set of scales, the Basic Interest Scales (BISs), provide more specific information about interests in 25 categories that are grouped within their respective GOTs. The third set of scales, the Occupational Scales (OSs), consist of 211 scales (102 paired occupations by gender and 7 gender-specific occupations) that compare respondents' interests to those of the criterion group employed in the designated occupations. The response format for the SII uses a 3-point Likert scale in which respondents indicate *like, indifferent*, or *dislike*. In addition, an optional 60-item, 6-scale *Skills Confidence Inventory* (Betz, Borgen, & Harmon, 1994) assesses self-efficacy for tasks associated with each of the six GOTs. The Skills Confidence Inventory uses a 5-point Likert scale in which respondents indicate their confidence in performing a particular task or school subject from 1 (*no confidence at all*) to 5 (*complete confidence*).

Harmon and Borgen (1995) have reported on the psychometric properties of the revised SII. The test–retest reliability of the six GOTs ranged from $r = .84$ to $r = .92$, and the alpha coefficients ranged from .90 to .94. The BISs have improved in both test–retest reliability and internal consistency. Test–retest reliability for college students and employed adults ranged from the low .80s to the low .90s. Cronbach's alpha coefficients ranged from .74 to .94. The OSs also demonstrated good reliability with internal consistency, ranging from the low .70s to the low .90s with different populations. Temporal stability for the occupational scales ranged from $r = .85$ to $r = .92$. The SII technical and applications guide (Harmon et al., 1994) provides detailed information regarding the validity of this version of the SII.

Concurrent and predictive validity have historically been robust for the predecessors of the SII; however, little validity information exists for the 1994 revision. Harmon et al. (1994) did provide information on the initial validity studies. Concurrent and construct validity for the GOTs in the 1994 version have been demonstrated by examining the rank order of the means on each of the occupations within

a GOT to show the accuracy of scale placement. Additional support for the concurrent and construct validity of the GOTs was demonstrated through the GOTs' ability to differentiate between college majors in theoretically expected ways. The concurrent validity of BISs has been supported by comparing the mean scores of people employed in different occupations. High scorers on these scales were employed in occupations relevant to a particular BIS, whereas lower scorers were employed in occupations unrelated to that BIS. Evidence for concurrent validity of the OSs has been demonstrated by the ability of the scales to discriminate between people employed in an occupation and the general reference sample. Further evidence of concurrent validity comes from the lack of overlap among occupational scales.

Studies that investigated both the internal consistency and temporal stability of the Skills Confidence Inventory have reported strong reliabilities. Test–retest reliabilities for the measure over a 3-week period ranged from $r = .83$ for the Realistic scale to $r = .87$ for the Social scale in a sample of 113 college students. Cronbach's alpha estimates of internal consistency ranged from .84 on the Realistic scale to .87 on the Enterprising scale in a college sample and from .84 to .88 in a sample of employed adults.

The Skills Confidence Inventory (Betz et al., 1994) has demonstrated both concurrent and construct validity (Betz, Borgen, Kaplan, & Harmon, 1998). Concurrent validity has been shown through the inventory's ability to discriminate between employed adults and college students in the theoretically expected direction, with those employed having higher skills confidence than that of college students. In addition, skills confidence scores discriminated among different occupations as theoretically expected. Construct validity has been established through correlations ranging between $r = .44$ and $r = .65$ for the Skills Confidence and GOT scores on the SII.

Revised UNISEX edition of the ACT Interest Inventory. The UNIACT-R consists of 90 items that describe work-related activities. Respondents answer *like, indifferent,* or *dislike* to the items. Stanine scores are reported for six scales that correspond to the RIASEC types: Technical (Realistic), Science (Investigative), Arts (Artistic), Social Service (Social), Business Contact (Enterprising), and Business Operations (Conventional). In addition, the inventory reports scores for two bipolar work task dimensions: Data/Ideas and People/Things. These two dimensions structure the World-of-Work Map that depicts how interests relate to 23 job families (Prediger, 1982).

Alpha coefficients for the UNIACT-R have been examined by age and sex (ACT, 1995). For the six interest scales, alpha coefficients ranged from .77 to .92 for males and from .83 to .92 for females. Split-half reliability for the Data/Ideas dimension was .83 for 8th-grade males, .87 for 10th-grade males, .88 for 8th-grade females, and $r = .92$ for 10th-grade females. Split-half reliability for the People/Things dimension was .67 for 8th-grade males, .74 for 10th-grade males, .71 for 8th-grade females, and .73 for 10th-grade females. Temporal stability at 4 years ranged from .47 to .60 for males and from .50 to .61 for females.

Numerous studies have supported the construct, convergent, discriminant, and criterion validity of the UNIACT-R (ACT, 1995). Correlations between the UNIACT-R scales occur in theoretically expected directions for high school and adult populations. In addition, scale configurations coincide with the RIASEC hexagon and the Data/Ideas and People/Things dimensions. Further support for construct validity has been shown through moderate correlations ($r = .49$ to $r = .65$ for males, $r = .45$ to $r = .64$ for females) between the interest scales and the activity experiences of 11th-grade students. Discriminant validity has been evidenced by lower correlations between dissimilar scales such as self-estimates of ability and tested abilities. Criterion validity has been shown through the hexagonal location of occupational groups and hit rates as well as through agreement among academic major, occupational choice, and RIASEC type.

Procedures

We approached this validation process by using Campbell and Fiske's (1959) data-analytic technique of the multitrait–multimethod matrix to compare the five methods of inventorying three sets of career constructs: vocational interests as indicated by RIASEC scales, occupational interests as indicated by occupational scales, and vocational self-efficacy as indicated by skill rating scales. Using the MTMM data-analytic technique, we entered the correlations between all of the variables on each of the inventories into matrices that organized the correlations into three groups: monotrait–heteromethod, heterotrait–monomethod, and heterotrait–heteromethod.

Convergent validity is indicated by the monotrait–heteromethod correlations. Convergent validity (Campbell & Fiske, 1959) requires that the same constructs, or constructs that share similar names, measured by different inventories produce substantial and positive correlations. Based on our literature review, we expected that occupational scales with similar names across inventories would have correlation coefficients in the upper .30s. The studies reviewed herein suggest that, for example, occupational scales such as the CISS Psychologist scale and the SII Psychologist scale should correlate about $r = .39$. Furthermore, we expected that scores for the RIASEC vocational interests across five inventories (e.g., SDS Realistic scale scores and SII Realistic scale scores) and for the RIASEC vocational self-efficacy scales between the CISS and SII (e.g., CISS Producing scale scores and SII Enterprising scale scores) would show higher correlations, about $r = .50$, based on our review of correlations between basic interest scales on different inventories and because, in tapping a broader domain, the RIASEC vocational scales should have less error than the occupational scales.

Discriminant validity refers to the degree to which measures of different constructs are unique. As noted by Campbell and Fiske (1959), definitions imply distinctions, and discriminant validity verifies these distinctions as an important part of the validation process. Scales can be invalidated if they correlate too highly to other scales that measure supposedly different traits. While convergent validity is shown by substantial and positive correlation between the same construct on

different inventories, discriminant validity is shown when these validity correlations exceed the correlations of the construct with other variables in the heterotrait blocks. This ratio provides evidence for relative validity in showing how greatly common variance specific to the variable exceeds shared method variance and error variance. The MTMM matrix can be used to evaluate three major criteria for discriminant validity by structuring three sets of comparisons: validity values to control values, validity values to method values, and method values to control values. The following paragraphs, in turn, explain how each of these three comparisons pertains to the three major criteria for discriminant validity.

First, discriminant validity is supported when validity values for a variable exceed the control values indicated by correlations obtained for that variable and different variables measured by different methods. In the MTMM matrix, the heterotrait–heteromethod block contains the correlations between different traits measured by different methods. Thus, we expected that correlations between dissimilar occupational scales and between different RIASEC scales measured across the four pertinent inventories would be less than the validity values for the matched scales. This means that a validity value should exceed the correlation coefficients lying in its row and column in the heterotrait–heteromethod blocks. For example, the SDS Realistic scale should correlate higher to the SII Realistic scale than to the SII Social scale, and the SII Psychologist scale should correlate higher to the CISS Psychologist scale than to the CISS Guidance Counselor scale.

Second, discriminant validity is supported when validity values exceed method variance. This means that correlations between the same variable measured by different methods should exceed correlations obtained for different variables measured by the same method. In the MTMM matrix, the heterotrait–monomethod block contains the correlations between different variables measured by the same method. When these heterotrait–monomethod correlations are compared to validity values (i.e., the monotrait–heteromethod correlations), they should be lower if the inventories are indeed measuring similar constructs independent of measurement method. For example, the correlation between the SDS Realistic scale and the SII Realistic scale should exceed the correlation between the SDS Social and Enterprising scales.

Whereas the second criterion for discriminant validity compares the magnitude of method correlations to validity correlations, the third criterion compares the magnitude and pattern of method correlations to the control correlations. Comparisons between these blocks provide important evidence regarding the potential interchangeability of trait and method variance. Systematic correlation between scale scores could be due to measurement features of the method and not the variable content representing the construct being measured. Accordingly, the third criterion for discriminant validity states that method variance should be absent or minimal.

We looked for the presence of a method factor by comparing both the magnitude and the pattern of correlations across the monomethod and heteromethod blocks. With regard to comparisons of magnitude, the values in the monomethod block include correlated measurement error, and any of this response set variance has the

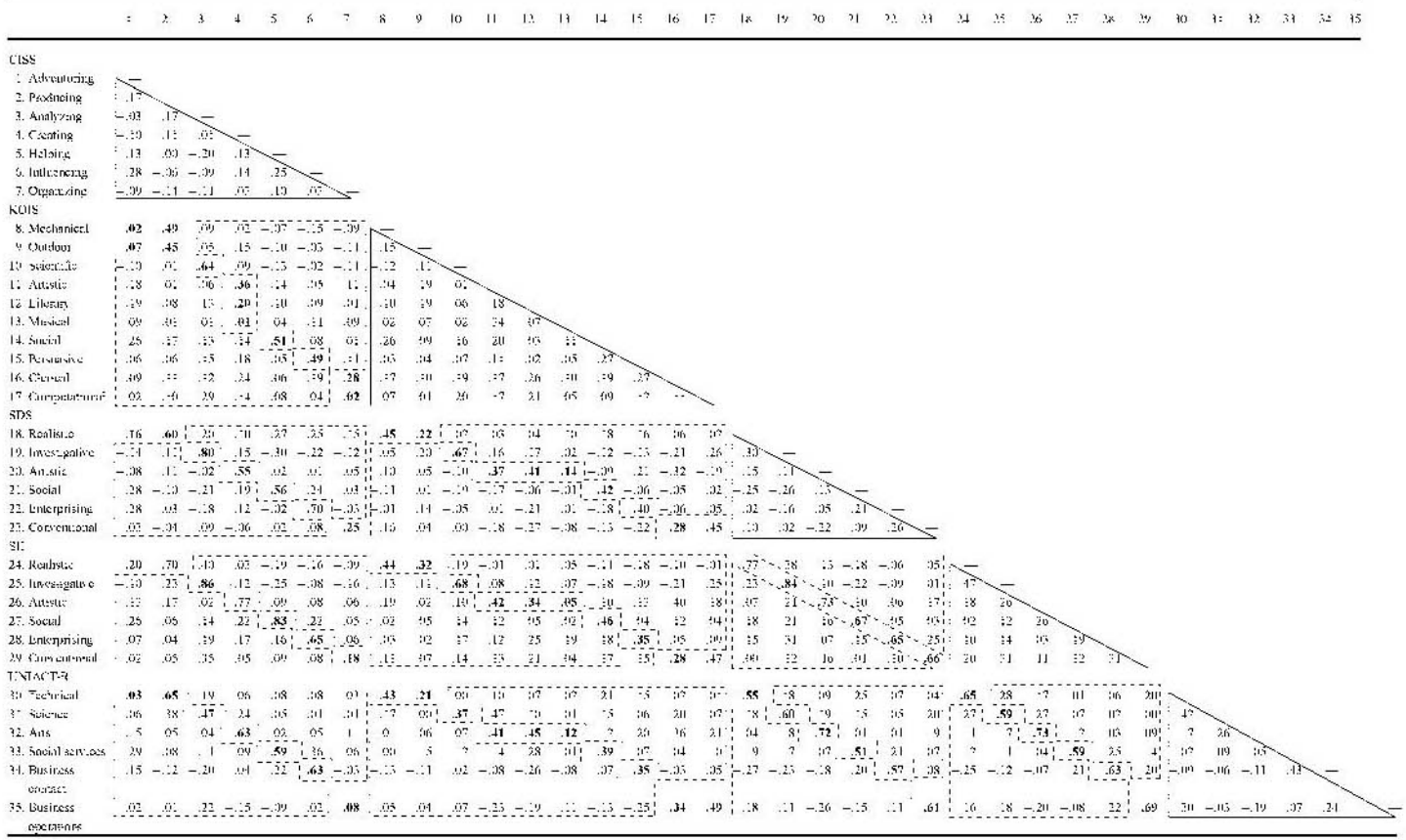
effect of increasing the degree of correlation in the validity diagonal. In contrast to the monomethod block, these measurement errors should be independent in the heteromethod blocks, producing lower correlations in the validity diagonal and heterotrait blocks. Accordingly, the strength of method variance was estimated by comparing the magnitude of "shared confounded irrelevancies" (Campbell & Fiske, 1959, p. 97) reflected by heterotrait–heteromethod correlations (error variance) to the heterotrait–monomethod correlations (response bias). In an MTMM matrix, variance that could be attributed to measurement method is shown by differences between the correlations for parallel constructs (heterotraits) measured by the same method (monomethod) and by different methods (heteromethod). For example, in the current study, a parallel comparison occurs for the correlation between the Realistic scale and the Social scale (heterotraits) on the SDS (monomethod) relative to the correlations between the Realistic scale on the SDS and the Social scales (parallel heterotraits) on each of the other four inventories (heteromethods). If the SDS Realistic and Social scales correlate higher than the SDS Realistic scale correlates to the Social scales on the other four inventories, then this difference estimates the strength of method variance.

A second estimate of the strength of method variance, distinct from differences in magnitude, is similarity in the patterns of correlation within the monomethod and heteromethod blocks. The third discriminant validity criterion also states that the pattern of variable correlations should be replicated across the monomethod block and each of the four heteromethod blocks. Because it is difficult to determine patterns by visual inspection of these five blocks, we calculated Kendall's coefficient of concordance to provide a quantitative estimate of pattern similarity (Maxwell, 1961). This calculation involved rank ordering the correlation coefficients in each of the five heterotrait blocks and then comparing the rank orders across the five blocks. The resulting coefficient of concordance (W) is the ratio of the obtained similarity to the maximum similarity possible when all rankings are identical. It ranges from .00 to 1.00. In the current study, W provides a measure of agreement among the rankings of the six RIASEC constructs by the five different inventories. Higher values of W indicate less method variance.

RESULTS AND DISCUSSION

The current study examined the convergent and discriminant validity of five methods for inventorying interests across three sets of scales that purport to measure vocational interests, occupational interests, and vocational self-efficacy. The full 35×35 MTMM matrix used in the analysis of vocational interest scales appears as Table 1. Summaries of the MTMM results for the five inventories' vocational interest scales appear in Tables 2 through 6. The summaries are divided into three sets of correlations: the monotrait–heteromethod (validity values), the heterotrait–monomethod (method values), and the heterotrait–heteromethod (error values). Within each set, we summarize the correlation values found in the convergent

TABLE 1
MTMM of the Vocational Interest Scales on the CISS, KOIS, SDS, SII, and UNIACT-R



diagonal, heterotrait–monomethod block, and heterotrait–heteromethod blocks by reporting the median value and the lowest and highest values. We report the medians rather than the means because, in this case, medians appear to provide more stable estimates of central tendency. Additional MTMM matrices were constructed for nine occupational scales (see Table 7) and the skill rating scales on the CISS, SDS, and SII (see Table 8). For each matrix, correlations equal to or greater than $r = .19$ were significant at the .05 level, and correlations equal to or greater than $r = .28$ were significant at the .01 level.

The results are presented in four major sections: (a) vocational interest scales, (b) occupational interest scales, (c) self-rating scales, and (d) items. Within each section, the results for each of the inventories are presented in the same sequence. We begin the sequence for each inventory with a consideration of convergent validity as indicated by the monotrait–heteromethod correlations between different measures of the same vocational interest constructs. After examining the evidence for convergent validity, we consider in turn each of the three criteria for discriminant validity. Relative to the first criterion for discriminant validity, we examine the MTMM matrices to determine whether the validity values exceed the control values indicated by the heterotrait–heteromethod correlations between each variable and different variables measured in different ways. Relative to the second discriminant validity criterion, we determine whether validity values exceed method values by comparing the monotrait–heteromethod correlations to the heterotrait–monomethod correlations between dissimilar variables measured by the same method. Relative to the third criterion for discriminant validity, we look for the presence of method variance by comparing the magnitude of correlations in the heterotrait–monomethod block to that in the heterotrait–heteromethod blocks. We also report a coefficient of concordance for the rankings of the six RIASEC constructs in the heterotrait blocks to indicate how well the pattern of variable correlations replicates within the five heterotrait blocks (i.e., one monomethod block and four heteromethod blocks). The presentation of these convergent and discriminant validity results for each set of scales on each inventory is immediately followed by a brief discussion of the meaning of the findings.

Vocational Interest Scales

Campbell Interest and Skills Survey

Table 2 summarizes the CISS validity, method, and control values. With respect to convergent validity, the median monotrait–heteromethod correlation coefficients for five of the seven CISS scales were above $r = .35$ and statistically significant at the .01 level. The median values on the five Orientation Scales ranged from $r = .46$ for Creating to $r = .72$ for Analyzing. The Organizing and Adventuring scales fared less well. The CISS Organizing scale achieved a median convergent validity coefficient of only $r = .18$; it did not correlate significantly to the SII Conventional, UNIACT-R Business Operations, or KOIS Computational scales, and it correlated only moderately to the SDS Conventional ($r = .25$, $p < .05$) and KOIS Clerical ($r = .28$, $p < .01$) scales. The Adventuring scale achieved a median

TABLE 2
MTMM Summary of the Vocational Interest Scales on the CISS

	Validity			Method			Error		
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum
Adventuring	.07	.02	.20*	.01	−.10	.28**	.00	−.17	.38**
Producing	.60	.17	.70**	.01	−.20*	.17	.00	−.19	.29**
Analyzing	.72	.47**	.86**	.01	−.20*	.17	.00	−.19**	.40**
Creating	.46	−.01	.77**	.01	−.10	.14	.01	−.24*	.24*
Helping	.58	.51**	.83**	.11	−.20*	.25**	−.01	−.30**	.22*
Influencing	.64	.49**	.70**	.01	−.09	.28**	.00	−.25**	.36**
Organizing	.18	−.02	.28*	.00	−.14	.10	.00	−.17	.11

*$p < .05$.
**$p < .01$.

convergent validity value of $r = .07$. The only significant correlation to scales that measure the Realistic construct on the other inventories was $r = .20$ ($p < .05$) to the SII Realistic scale.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the CISS validity values exceeded the control values. The control correlations ranged from $r = -.30$ between the CISS Helping scale and the SDS Investigative scale to $r = .40$ between the CISS Analyzing scale and the SII Realistic scale. Of the 163 heterotrait–heteromethod correlations, 31 were statistically significant. Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 6.58$, $p < .001$). This means that the CISS scales showed good evidence of discriminant validity. This strong pattern, however, should not obscure the serious problem indicated by the instance in which the error variance exceeded the shared common variance. This problem occurred with the Adventuring scale because it had significant positive correlations to three scales that measure the Social construct (UNIACT-R Social Service $r = .29$, KOIS Social $r = .26$, and SII Social $r = .26$) and to two scales that measure the Enterprising construct (CISS Influencing $r = .28$ and SDS Enterprising $r = .28$) as well as a significant negative correlation to the KOIS Literary scale ($r = -.19$). These six control values were higher than the validity values for the Adventuring scale. This indicated that the Adventuring scale shared more variance with seemingly unrelated scales than it did with the Realistic scales, which the CISS authors purported Adventuring to be a facet. Thus, the Adventuring scale failed to meet a minimum requirement for validity. Other positive moderate correlations were observed for CISS Analyzing to SII Realistic ($r = .40$), CISS Producing to UNIACT-R Science ($r = .38$), and CISS Influencing to UNIACT-R Social Service ($r = .36$). Moderate negative correlations were observed in the correlations of CISS Helping to SDS Investigative ($r = -.30$) and to SDS Realistic ($r = -.27$) as well as in the correlations of CISS Influencing to SDS Realistic ($r = -.25$) and to SDS Investigative ($r = -.22$).

Relative to the second discriminant validity criterion, we determined whether the CISS validity values exceeded its method values. The method correlations ranged from $r = -.20$ between the CISS Analyzing and Helping scales to $r = .28$ between the CISS Adventuring and Influencing scales. Only three of the 21 heterotrait–monomethod correlations for the CISS were statistically significant: Adventuring to Influencing scales ($r = .28$), Analyzing to Helping scales ($r = -.20$), and Helping to Influencing scales ($r = .25$). The validity values in all cases exceeded the method values, although for the Organizing and Adventuring scales the validity values sometimes only slightly exceeded the method correlations (e.g., the validity value for CISS Organizing to UNIACT-R Business was $r = .08$ compared to method values of $r = .03$ to CISS Influencing and $r = .03$ to CISS Creating). Compared to the values in the validity diagonal, proportionately fewer significant method correlations were observed ($z = 4.41$, $p < .001$). This finding meets the second discriminant validity criterion that shared common variance should exceed response bias variance.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the magnitude and pattern of correlations. With regard to magnitude, the median CISS correlation coefficients were essentially zero for both the method and control values, with a single exception being that the method correlation for the Helping scale was $r = .11$ compared to a control correlation of $r = .01$. Relative to the pattern of correlations, the coefficient of concordance for the rankings of the six RIASEC constructs in the CISS heterotrait blocks was $W = .55$ ($p < .001$), indicating that the rankings were not random and showing a moderate degree of pattern similarity. Taken together, the relative magnitude and pattern similarity indicators for the presence of response bias suggested that the CISS contains minimal method variance and therefore meets the third criterion for discriminant validity.

Discussion of CISS results. Five of the seven CISS Orientation Scales showed strong convergent validity coefficients, demonstrating validity as measures of Holland's RIASEC constructs. However, the convergent validity coefficients for the Adventuring and Organizing scales did not support their validity as measures of RIASEC constructs. The CISS Adventuring scale failed to converge with other Realistic scales, and it correlated higher to Social and Enterprising scales on the other inventories. This finding suggests that the Adventuring scale may need to be redesigned or redefined. In constructing the CISS, Campbell split Holland's Realistic type into the Producing orientation, which includes mechanical, construction, and farming activities, and the Adventuring orientation, which includes military, police, and athletic activities (Campbell et al., 1992). In the current study, the validity correlations for the Producing scale met expectations, but those for the Adventuring scale did not. Adventuring had a low yet statistically significant correlation of $r = .28$ to both the SDS Enterprising scale and the CISS Influencing scale. This finding might be explained by the fact that in Campbell's model the most similar scales to Adventuring are Producing on one side and Influencing on the other side. However, in that model, Adventuring is opposite to Helping,

indicating that they are the most dissimilar. This is at odds with the Adventuring scale having its highest correlation on each of the other four inventories to the Social scale. Perhaps the military and police interests in the Adventuring orientation of the CISS include a component of social service, making the use of physical skill a contribution to the community.

The other CISS orientation scale that produced unexpected results was Organizing. In Campbell's model, the Organizing scale differs from Holland's Conventional type in leaning more toward management and financial services than toward office and clerical work. The Organizing scale did not converge with any scales on the SII and UNIACT-R, and it only weakly converged with the KOIS Clerical scale ($r = .28$) and the SDS Conventional scale ($r = .25$). It did not correlate significantly to any other scales in the MTMM matrix. This probably means that the Organizing scale succeeds in measuring management and financial service interests and that these interests are more distinct from clerical and office interests, and any other vocational interests, than we may have thought.

With regard to discriminant validity, the CISS scales met the criterion of validity values exceeding both control and method values and in demonstrating minimal method variance, with the single exception of the Adventuring scale, for which several control values and one method value (for the Social and Enterprising scales) exceeded the CISS validity values.

Kuder Occupational Interest Survey

Table 3 summarizes the KOIS validity, method, and control values. With regard to convergent validity, the median monotrait–heteromethod correlations for the 10 scales ranged from $r = .09$ for the Musical scale to $r = .66$ for the Scientific scale. Of the 10 median convergent validity coefficients for the KOIS, 7 were significant at the .01 level. Two of the three remaining scales had median convergent validity coefficients that were significant at the .05 level: Outdoor ($r = .22$) and Literary ($r = .38$). The Musical scale did not have statistically significant

TABLE 3
MTMM Summary of the Vocational Interest Scales on the KOIS

	Validity			Method			Error		
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum
Mechanical	.44	.02	.49**	.00	−.26**	.15	.00	−.19	.19*
Outdoor	.22	.07	.45**	.00	−.19*	.15	.00	−.15	.37**
Scientific	.66	.37**	.68**	.00	−.19*	.20*	.00	−.19*	.19*
Artistic	.39	.36**	.42**	−.14	−.34**	.18	−.01	−.23*	.16
Literary	.38	.20*	.45**	.01	−.26**	.18	−.01	−.28**	.19*
Musical	.09	−.01	.14	.00	−.34**	.11	−.01	−.19	.11
Social Service	.44	.39**	.51**	−.16	−.27**	.11	−.13	−.21*	.26**
Persuasive	.38	.35**	.49**	.01	−.27**	.11	−.01	−.25**	.21*
Clerical	.28	.28**	.34**	−.17	−.27**	.11	−.01	−.40**	.05
Computational	.46	−.02	.49**	.00	−.21*	.20*	.00	−.21**	.29**

*$p < .05$.
**$p < .01$.

correlations and thus did not converge with scales that measure the construct of artistic vocational interests on the other four inventories.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the KOIS validity values exceeded the control values. The control correlations ranged from $r = -.40$ between KOIS Clerical and SII Artistic to $r = .47$ between the KOIS Artistic and UNIACT-R Science. Of the 208 heterotrait–heteromethod correlations, 31 were statistically significant. Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 9.18$, $p < .001$). This means that generally the KOIS scales showed discriminant validity. The significant control correlations were low and negative, with two exceptions. The KOIS Outdoor scale and the SII Investigative scale correlated $r = .20$, and the KOIS Clerical scale and the SII Artistic scale correlated $r = -.40$.

Relative to the second discriminant validity criterion, we determined whether the KOIS validity values exceeded its method values. The method correlations ranged from $r = -.34$ between the KOIS Artistic and Musical scales to $r = .28$ between the KOIS Scientific and Computational scales. Of the 45 heterotrait–monomethod correlations, 12 were significant. In general, the significant method correlations were low and negative. The only notable exceptions to this were the correlation between the Musical and Artistic scales ($r = -.34$) and the correlations of the Persuasive scale to the Clerical ($r = -.27$) and Social ($r = .27$) scales. In every case, the convergent values exceeded the method values. Compared to the monotrait–heteromethod correlations, proportionately fewer significant heterotrait–monomethod correlations were observed ($z = 5.30$, $p < .001$). These findings meet the second criterion for discriminant validity that shared common variance should exceed response bias variance.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the magnitude and pattern of correlations. With regard to magnitude, the median CISS correlation coefficients in both the method and control blocks were essentially zero. Two scales did show some variance attributable to features of the inventory rather than the traits being measured. The Clerical scale had a median method coefficient of $r = -.17$ compared to a median control coefficient of $r = -.01$, and the Artistic scale had a median method coefficient of $r = -.14$ compared to a median control coefficient of $r = -01$. Relative to the pattern of variable correlations, the coefficient of concordance for the rankings of the six RIASEC constructs in the KOIS heterotrait blocks was $W = .45$ ($p < .001$), indicating that the rankings were not random and showing a moderate degree of pattern similarity. Taken together, the relative magnitude and pattern similarity indicators for the presence of response bias suggested that the KOIS contains minimal method variance and therefore meets the third criterion for discriminant validity.

Discussion of KOIS results. Only one of the KOIS scales met the expectation of converging with relevant RIASEC scales on other inventories with a validity value at or above $r = .50$. Moreover, the KOIS as a measure of RIASEC constructs had the

lowest validity values among the five inventories. This could be attributed to the fact that the KOIS uses measurement methods that differ substantially from those of the other inventories. Furthermore, Kuder designed the KOIS vocational interest scales long before Holland's typology became prominent and influenced the construction of the other four inventories. Also, two RIASEC constructs, Conventional and Realistic, are measured by two different KOIS scales. Although Mechanical and Outdoor converge, note for interpretive purposes that the KOIS Mechanical scale had stronger correlations to the Realistic scales on the other four inventories. The Clerical and Computational scales converged, but Computational had stronger correlations to other Conventional scales and showed a relation to Investigative interests that the Clerical scale lacked. The KOIS measures the Artistic type with three separate scales: Artistic, Literary, and Musical. The KOIS Musical scale failed to converge with the Artistic scales on the other four inventories, and its only significant correlation was a method correlation of $r = -.34$ to the KOIS Artistic scale. Clearly, the Musical scale lacks construct validity as a measure of Holland's Artistic construct. The KOIS Literary and Artistic scales both correlated about $r = .40$ to the Artistic scales in the other four inventories. It appears that the Creating and Artistic scales on the CISS and SII lean a little more toward art, whereas the Artistic scales on the SDS and UNIACT-R lean a little more toward writing. With regard to discriminant validity, the KOIS scales met the criterion of validity values exceeding control and method values and in demonstrating minimal method variance.

Self-Directed Search

Table 4 summarizes the SDS validity, method, and control values. With regard to convergent validity, the median monotrait–heteromethod correlations ranged from $r = .45$ for the Conventional scale to $r = .82$ for the Investigative scale. All of the median convergent validity values were statistically significant at the .01 level. SDS scales tended to correlate the highest with their corresponding SII scales.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the SDS validity values exceeded the control values. The control correlations

TABLE 4
MTMM Summary of the Vocational Interest Scales on the SDS

	Validity			Method			Error		
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum
Realistic	.55	.16	.77**	.01	−.25**	.30**	.00	−.27**	.23*
Investigative	.82	.60**	.86**	.00	−.26**	.30**	.01	−.31**	.38**
Artistic	.48	.14	.73**	.13	−.22*	.15	.00	−.32**	.21*
Social	.54	.42**	.66**	.01	−.26**	.24	−.01	−.25**	.29**
Enterprising	.61	.40**	.70**	.01	−.16	.26**	.01	−.21*	.28**
Conventional	.45	.25*	.66**	.01	−.22*	.26**	.00	−.27**	.25**

*$p < .05$.
**$p < .01$.

ranged from $r = -.32$ between the SDS Artistic scale and the KOIS Clerical scale to $r = .38$ between the SDS Investigative scale and the SII Realistic scale. Of the 145 heterotrait–heteromethod correlations, 31 were statistically significant. Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 7.48$, $p < .001$). As with other inventories, the heterotrait–heteromethod correlations were generally low. The highest positive correlations occurred between adjacent types on different inventories; for example, the SDS Investigative scale and the SII Realistic scale correlated $r = .38$. The two highest negative correlations occurred between opposite types on different inventories; the SDS Artistic scale and the KOIS Clerical scale correlated $r = -.32$, and the SDS Investigative scale and the SII Enterprising scale correlated $r = -.31$.

Relative to the second discriminant validity criterion, we determined whether the SDS validity values exceeded its method values. The method correlations ranged from $r = -.26$ between the SDS Investigative and Social scales to $r = .30$ between the SDS Realistic and Investigative scales. Of the 15 heterotrait–monomethod correlations, 6 were significant. Half of the values were positive and half were negative. The three positive correlations all occurred between adjacent types: Realistic to Investigative ($r = .30$), Enterprising to Social ($r = .24$), and Enterprising to Conventional ($r = .26$). Two of the three negative correlations occurred between opposite types: Social to Realistic ($r = -.25$) and Artistic to Conventional ($r = -.22$). The sixth statistically significant correlation occurred between Social and Investigative ($r = -.26$). Compared to the monotrait–heteromethod correlations, proportionately fewer significant heterotrait–monomethod correlations were observed ($z = 3.86$, $p < .001$). This finding meets the second discriminant criterion that shared common variance should exceed response bias variance.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance by comparing the magnitude and pattern of correlations. With regard to magnitude, the SDS median method values were the same as or smaller than the median control values, with the single exception of the Artistic scale, for which the median method correlation was $r = .13$ compared to $r = .00$ for the median control correlation. Relative to the pattern of correlations, the coefficient of concordance for the rankings of the six RIASEC constructs in the SDS heterotrait blocks was $W = .60$ ($p < .001$), indicating that the rankings were not random and showing a moderate degree of pattern similarity. Taken together, the relative magnitude and pattern similarity indicators for the presence of response bias suggested that the SDS contains minimal method variance and therefore meets the third criterion for discriminant validity.

Discussion of SDS results. Scales for all six RIASEC constructs had median validity values at or above $r = .45$. They converged most strongly to the RIASEC scales in the SII, with a median validity coefficient of $r = .70$, and weakest to the KOIS, with a median validity coefficient of $r = .41$. Convergence with the UNIACT-R was $r = .58$ and with the CISS was $r = .58$. This finding coincides with the widely held belief that the best measures of the RIASEC types as defined

TABLE 5
MTMM Summary of the Vocational Interest Scales on the SII

	Validity			Method			Error		
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum
Realistic	.55	.20*	.77**	.18	−.10	.47**	.00	−.25**	.40**
Investigative	.76	.60**	.86**	.26	−.14	.47**	.01	−.25**	.28**
Artistic	.55	.05	.77**	.18	−.11	.26**	.01	−.40**	.27**
Social	.59	.46**	.83**	.12	−.12	.26**	.00	−.21*	.26**
Enterprising	.64	.35**	.65**	.00	−.14	.31**	.00	−.31**	.25**
Conventional	.47	.18	.69**	.20	−.11	.31**	.00	−.21*	.35**

*$p < .05$.
**$p < .01$.

by Holland are the SDS and the SII. With regard to discriminant validity, the SDS scales met the criterion of validity values exceeding control and method values and in demonstrating minimal method variance.

Strong Interest Inventory

Table 5 summarizes the SII validity, method, and control values. With regard to convergent validity, the median monotrait–heteromethod correlations ranged from $r = .47$ for the Conventional scale to $r = .76$ for the Investigative scale. All of the median convergent validity values were statistically significant at the .01 level. The SII Conventional scale failed to converge significantly with the CISS Organizing scale ($r = .18$), and the SII Artistic scale failed to converge with the KOIS Musical scale ($r = .05$).

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the SII validity values exceeded the control values. The control correlations ranged from $r = -.40$ between the SII Artistic scale and the KOIS Clerical scale to $r = .40$ between the SII Realistic scale and the CISS Analyzing scale. Of the 145 heterotrait–heteromethod correlations, 33 were statistically significant. Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 7.28$, $p < .001$). This means that the SII showed good discriminant validity.

Relative to the second discriminant validity criterion, we determined whether the SII validity values exceeded its method values. The method correlations ranged from $r = -.14$ between the SII Investigative and Enterprising scales to $r = .47$ between the SII Realistic and Investigative scales. Of the 15 heterotrait–monomethod correlations, 7 were statistically significant. The highest significant correlations occurred between adjacent scales on the RIASEC hexagon. Compared to the monotrait–heteromethod correlations, proportionately fewer significant heterotrait–monomethod correlations were observed ($z = 3.48$, $p < .001$). This finding meets the second discriminant validity criterion, which states that shared common variance should exceed response bias variance.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the magnitude and pattern of correlations. With regard to magnitude, the SII median method and control correlations for the Enterprising scale both were zero. However, for the other five scales, the median control correlations were much smaller than the median method correlations. The biggest discrepancy was between the Investigative scale's method value of $r = .28$ and its control value of $r = .01$. Relative to the pattern of correlations, the coefficient of concordance for the rankings of the six RIASEC constructs in the SII heterotrait blocks was $W = .61$ ($p < .001$), indicating that the rankings were not random and showing a moderate degree of pattern similarity. Taken together, the relative magnitude and pattern similarity indicators for the presence of response bias suggested that the SII contains some variance attributable to measurement method yet stills meets the third criterion for discriminant validity.

Discussion of SII results. For the SII, all six RIASEC scales had validity values at or above $r = .47$. The most notable anomaly was that the SII Conventional scale did not correlate significantly with the CISS Organizing scale ($r = .18$). We have already explained that the CISS Organizing scale is unique in measuring management and financial services rather than office and clerical work. The only other anomaly was that the SII Artistic scale failed to converge with the KOIS Musical scale. With regard to discriminant validity, the SII scales strongly met the criterion of validity values exceeding control and method values. However, the SII did demonstrate some method variance, with median method values generally exceeding control values by about .20. In studying the pattern of method correlations, however, it looks as though the shared variance can be accounted for by correlations between types adjacent on the hexagon, which fits the RIASEC theory.

Revised UNIACT

Table 6 summarizes the UNIACT-R validity, method, and control values. With regard to convergent validity, the median validity values ranged from $r = .49$ for both the Technical and Business Operations scales to $r = .60$ for the Business

TABLE 6
MTMM Summary of the Vocational Interest Scales on the UNIACT-R

	Validity			Method			Error		
	Median	Minimum	Maximum	Median	Minimum	Maximum	Median	Minimum	Maximum
Technical	.49	.03	.65**	.16	−.09	.42**	.00	−.25**	.28**
Science	.53	.37**	.60**	.00	−.09	.42**	.00	−.20*	.38**
Arts	.54	.12	.73**	.01	−.19*	.26**	.00	−.36**	.20
Social service	.55	.39**	.59**	.01	−.09	.43**	.00	−.28**	.36**
Business contact	.60	.35**	.63**	.01	−.19*	.30**	−.01	−.27**	.22*
Business operations	.49	.07	.69**	.01	−.11	.43**	.00	−.26**	.22*

*$p < .05$.
**$p < .01$.

Contact scale. All of the median convergent validity correlations were statistically significant at the .01 level. The UNIACT-R correlated highest to the SII for Investigative, Artistic, Enterprising, and Conventional vocational interests and correlated highest to the CISS for Producing and Creating vocational interests.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the UNIACT-R validity values exceeded the control values. The control correlations ranged from $r = -.36$ between the UNIACT-R Arts scale and the KOIS Clerical scale to $r = .38$ between the UNIACT-R Science scale and the CISS Producing scale. Of the 145 heterotrait–heteromethod correlations, 41 were significant. Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 6.20$, $p < .001$). This means that the UNIACT-R showed good evidence of discriminant validity.

Relative to the second discriminant validity criterion, we determined whether the UNIACT-R validity values exceeded its method values. The method correlations ranged from $r = -.19$ between the UNIACT-R Arts and Business Operations scales to $r = .43$ between the UNIACT-R Social Service and Business Contact scales. Of the 15 heterotrait–monomethod correlations, 6 were significant. These correlations occurred between adjacent scales (positive correlation) or opposite (negative correlation) in the RIASEC hexagon. The highest two of these six correlations occurred between the Technical and Science scales ($r = .42$) and the Social Service and Business Contact scales ($r = .43$). The strongest negative correlation occurred between the Arts and Business Operations scales ($r = -.19$). Compared to the monotrait–heteromethod correlations, proportionately fewer statistically significant heterotrait–monomethod correlations were observed ($z = 3.51$, $p < .001$). This finding meets the second discriminant validity criterion, which states that shared common variance should exceed response bias variance.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the magnitude and pattern of correlations. With regard to magnitude, the UNIACT-R median correlation coefficients were essentially $r = .00$ for both the method and discriminant values, with the single exception being that the median method correlation for the Technical scale was $r = .16$ compared to a control correlation of zero. With regard to the pattern of correlations, the coefficient of concordance for the rankings of the six RIASEC constructs in the UNIACT-R heterotrait blocks was $W = .61$ ($p < .001$), indicating that the rankings were not random and showing a moderate degree of pattern similarity. Taken together, the relative magnitude and pattern similarity indicators for the presence of response bias suggested that the UNIACT-R contains minimal method variance and therefore meets the third criterion for discriminant validity.

Discussion of UNIACT-R results. All six RIASEC scales had median validity values at or above $r = .49$. They showed a strong and highly similar pattern of convergence with the SDS and SII. They showed the same results for the CISS, except for the failure of the UNIACT-R Business Operations scale to converge with the CISS Organizing scale, again probably because one measures management

and financial services and the other measures office and clerical activities. The UNIACT-R showed consistent yet generally weaker convergence with the KOIS. With regard to discriminant validity, the UNIACT-R scales met the criterion of validity values exceeding control and method values and of demonstrating minimal method variance.

Occupational Scales

We computed an MTMM matrix (available from the authors) for all of the matched occupational scales on the three inventories. We identified 22 matched occupational scales on the CISS, KOIS, and SII for a total of 66 scales; 14 matched occupational scales on the CISS and SII for a total of 28 scales; and 1 matched scale on the KOIS and SII for a total of 2 scales. In examining the results for all of these 96 scales, we concentrated on coefficients in the diagonals of the 96×96 MTMM. These convergent validity values can be compared directly to the correlations between matched occupational scales reported in prior studies. The median correlation for the 66 coefficients representing the 22 matched scales on the CISS, KOIS, and SII was $r = .42$. The median for the 22 correlation coefficients between the CISS and KOIS was $r = .32$; between the KOIS and SII, it was $r = .39$; and between the CISS and SII, it was $r = .63$. The correlations between the CISS and SII were significantly higher than the correlations of the KOIS scales to the CISS scales and to the SII scales. There were 14 more matched scales between the CISS and SII. Their median correlation coefficient was $r = .45$, making a median correlation coefficient of $r = .55$ for 36 matched scales between the CISS and SII.

Across the three inventories, median correlations for the 22 matched occupational scales ranged from $r = .62$ for Chemist to $r = .15$ for Electrician. In general, the higher correlation coefficients ($r > .50$) occurred for Investigative and Artistic occupations that require advanced education such as Chemist ($r = .62$), Architect ($r = .63$), Psychologist ($r = .54$), Veterinarian ($r = .48$), Computer Programmer ($r = .53$), and Physician ($r = .41$), whereas the lower correlation coefficients ($r < .39$) occurred for Realistic and Conventional occupations that require less education such as Electrician ($r = .15$), Police Officer ($r = .36$), Bookkeeper ($r = .17$), Secretary ($r = .21$), Accountant ($r = .31$), and Carpenter ($r = .41$). In general, the middle correlation coefficients ($r = .39$ to $r = .49$) occurred for Enterprising occupations such as Attorney ($r = .43$), Human Resource Director ($r = .34$), Real Estate Salesperson ($r = .38$), and Bank Manager ($r = .41$) as well as for Social occupations such as Social Worker ($r = .42$), Librarian ($r = .45$), Guidance Counselor ($r = .31$), Math/Science Teacher ($r = .29$), Nurse ($r = .43$), and Religious Leader ($r = .32$). Of course, this grouping was not perfect, yet the trend was sufficiently clear to confirm the expectation for higher correlations between scales for occupations in which the workers are more homogeneous and have advanced degrees.

An analysis of discriminant validity for the 94 scales requires more space that even a monograph permits. Nevertheless, because we were interested in some indication of the discriminant validity of the occupational scales, we made a compromise. We constructed an MTMM matrix (see Table 7) for the three matched

TABLE 7
MTMM for Similarly and Same-Named Occupational Scales of the CISS, KOIS, and SII

	CISS			KOIS			SII		
	Guidance counselor	Human resource director	Psychologist	Guidance counselor	Personnel manager	Psychologist	High school counselor	Human resource director	Psychologist
CISS									
Guidance Counselor	—								
Human Resource Director	.06	—							
Psychologist	.18	−.18	—						
KOIS									
Guidance Counselor	**.31**	.12	.27	—					
Personnel Manager	.27	**.34**	.11	.70	—				
Psychologist	.06	.11	**.54**	.74	.63	—			
SII									
High School Counselor	**.67**	.40	−.02	**.25**	.25	.05	—		
Human Resource Director	.40	**.58**	.50	.11	**.30**	.21	.65	—	
Psychologist	−.17	.05	**.55**	.12	−.07	**.54**	.09	−.38	—

occupational scales on the CISS, KOIS, and SII that we thought were closest in job duties to career counselors and vocational psychologists: the same-named "Psychologist" scales on the three inventories, the similarly named "Human Resource Director" scales on the CISS and SII and "Personnel Manager" scale on the KOIS, and the similarly named "Guidance Counselor" scales on the CISS and KOIS and "High School Counselor" scale on the SII.

In examining the MTMM for these nine scales, we first considered convergent validity. The nine monotrait–heteromethod correlations (which appear in bold in Table 7) were each statistically significant at the .01 level. The monotrait–heteromethod correlations generally demonstrated convergence. They ranged from a high of $r = .67$ between the CISS Guidance Counselor scale and the SII High School Counselor scale to a low of $r = .25$ between the KOIS Guidance Counselor scale and the SII High School Counselor scale. For the Human Resource Director scales and Personnel Manager scale, correlations were higher between the CISS and SII ($r = .58$), whereas the KOIS Personnel Manager scale correlated $r = .34$ to the CISS Human Resource Director scale and $r = .30$ to the SII Human Resource Director scale. For the Psychologist scales, the correlation between the CISS and SII was $r = .55$, and the KOIS correlated $r = .54$ to both the CISS and SII. For the Guidance Counselor scales, the highest correlation occurred between the CISS and SII ($r = .67$), followed by the KOIS and CISS ($r = .31$) and then the KOIS and SII ($r = .25$). The median correlation for the nine matched scales was $r = .45$ (Human Resource Director $r = .41$, Psychologist $r = .54$, and Guidance Counselor $r = .33$).

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the CISS validity values exceeded the control values. The control correlations ranged from $r = -.50$ to $r = .40$. Proportionately there was a significant difference ($z = 2.07$, $p = .04$) between the monotrait–heteromethod correlations (validity values) and the heterotrait–heteromethod correlations (control values), suggesting that error variance does not contribute to the CISS occupational scale scores. None of the error correlations exceeded the relevant convergent validity coefficients. However, the KOIS Guidance Counselor validity coefficient to SII High School Counselor ($r = .25$) was the same as the correlation between the SII High School Counselor scale and the KOIS Personnel Manager scale. Of the 12 heterotrait–heteromethod correlations, 4 were significant. Two of the significant error correlations occurred with the CISS Guidance Counselor scale, and one each occurred with the CISS Psychologist scale and the CISS Human Resource Director scale. For the KOIS, there was a significant difference in the proportions compared to the convergent validity coefficients ($z = 2.13$, $p < .03$), suggesting that error variance does not contribute to the occupational scale scores on the KOIS. None of the error correlations exceeded the relevant convergent validity coefficients. Of the 12 error correlations, 3 were significant. Two of the error correlations occurred with the KOIS Personnel Manager scale and one with the Guidance Counselor scale. For the SII, proportionately there was a significant difference in the number of significant

correlations compared to the validity coefficients ($z = 12.13$, $p = .03$), suggesting that error variance does not contribute to the occupational scale scores on the SII. None of these error correlations exceeded the relevant convergent validity coefficients, but 3 of the 12 error correlations were significant. Two of these significant error correlations occurred with the High School Counselor scale, and the other significant error correlation occurred with the Human Resource Director scale.

Relative to the second discriminant validity criterion, we determined whether the validity values exceeded the method values. For the CISS, the heterotrait–monomethod correlations were low between the Human Resource Director and the Psychologist scales, ranging from $r = -.18$ to $r = .18$. For the CISS, proportionately there was a significant difference between the monotrait–heteromethod correlations and the heterotrait–monomethod correlations ($z = 2.45$, $p = .01$), indicating that method variance did not seem to increase the CISS convergent validity correlations. For the KOIS, the three heterotrait–monomethod correlations were each higher than the convergent validity values, indicating that method variance strongly affects the KOIS occupational scale scores. The correlation between the Personnel Manager and Psychologist scales was $r = .63$; for the Personnel Manager and Guidance Counselor scales, it was $r = .70$; and between the Psychologist and Guidance Counselor scales, it was $r = .74$. For the KOIS, proportionately there was not a significant difference between the number of significant monotrait–heteromethod and heterotrait–monomethod correlations because all of the method correlations and all of the convergent correlations were significant, indicating that method variance may be increasing the KOIS convergent validity correlations. For the SII, two of the three heterotrait–monomethod correlations were significant; the Human Resource Director scale correlated $r = -.38$ to the Psychologist scale and $r = .65$ to the High School Counselor scale. The correlation between the Psychologist and High School Counselor scales was not significant ($r = .09$). For the SII, proportionately there was not a significant difference between the monotrait–heteromethod correlations and the heterotrait–monomethod correlations ($z = 1.10$, $p = .27$), indicating that method variance may be increasing the SII convergent validity coefficients. In particular, method variance appears to affect the SII High School Counselor scale given that it correlates nearly the same to the SII Human Resource Director scale ($r = .65$) and the CISS Guidance Counselor scale ($r = .67$) but much lower to the KOIS Guidance Counselor scale ($r = .25$).

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the occupational scales by comparing the magnitude and pattern of correlations. With regard to magnitude, the median control correlation was $r = .11$. The median method correlations were $r = .06$ for the CISS, $r = .70$ for the KOIS, and $r = .09$ for the SII. The relative magnitude and pattern similarity indicators strongly suggest the presence of significant method variance in the KOIS but not in the CISS and SII, possibly reflecting the uniqueness of the KOIS scale construction relative to the more similarly constructed CISS and SII. With regard to the pattern of correlations, the coefficients of concordance were $W = .36$ ($p < .17$) for the CISS, $W = .52$ ($p = .07$) for the KOIS, and $W = .76$ ($p = .02$) for the SII.

These results are less stable than the other coefficients of concordance reported earlier in the current article because, in using only three scales, there was only 5 degrees of freedom and minor differences strongly affected the value of W. This does not diminish the strength of the findings for the SII, but it mitigates the interpretation of the lower values of W for the KOIS and CISS.

Discussion of results for occupational scales. Convergent validity was lower for occupational scales than for the vocational interest scales but was still in line with expectations. The median correlation for the 22 matched scales on the three inventories was $r = .42$, with the highest correlation being $r = .86$. This coincided with our hypothesis that occupational scales with similar names across inventories would have correlation coefficients in the upper .30s. We made this prediction because, in eight prior studies that compared the KOIS and SVIB, the median correlation between matched scales was $r = .34$, with a mode of $r = .39$. In the current study, the correlation between the KOIS and SII of $r = .38$ resembles the findings of three studies in which the KOIS and SVIB correlated $r = .39$ and another study in which they correlated $r = .37$. We should note that the highest correlation, $r = .86$, exceeded the upper limit of $r = .70$ deduced by Kuder (1969) in reasoning about the possible range of correlation resulting from inventories that use different methods to deal with common variance. This correlation of $r = .86$, however, does not exceed the upper limit of possible correlation based on scale internal consistency coefficients. This upper limit is estimated as the square root of the internal consistency coefficient. For example, an interest scale with an internal consistency coefficient of .81 may correlate to other scales as high as $r = .90$.

The median correlation of $r = .42$ for all three inventories, however, masks some important differences. The median correlation of $r = .55$ obtained between the 22 matched scales on the CISS and SII was substantially higher than the median correlation of the KOIS to the CISS ($r = .32$) and to the SII ($r = .39$). Contrasting the substantially higher correlations of the SII to the CISS suggests the possibility that research on matched scales has produced lower than expected correlation coefficients because of the substantial differences in how Kuder (1977) and Strong (1943) constructed their inventories, including how scale scores are calculated. KOIS scores are correlation coefficients that indicate the degree of similarity between a respondent's interest pattern and the interest pattern obtained both by a specific occupational group and by people in general, whereas SII scores are the sum of like responses for a set of items that differentiate an occupational criterion group from people in general. According to Kuder (1969), this difference lowers the upper limit for interscale correlation from about $r = .75$, based on scale reliabilities alone, to about $r = .45$ when adjusted for differences in the use of core common interests. The correlation of $r = .38$ between the KOIS and SII obtained in the current study is about as high as can the expected if one agrees with Kuder's reasoning.

The current results also suggest the possibility that scales for occupations in which workers are more homogeneous and have advanced degrees correlate more strongly. The homogeneity of a criterion group refers to the extent to which

members of that group express the same preferences by marking the same responses on an interest inventory. This may, in part, explain the current findings because, at least on the KOIS, the occupational groups differ in homogeneity, and the higher the homogeneity, the higher the possible score. Kuder and Zytowski (1991, p. 26) explained that when comparing scale scores for the KOIS and other inventories, optimal similarity requires that the scores have the same range from occupation to occupation.

The portrait of validity in prior studies relies exclusively on coefficients in the validity diagonal in isolation from the total MTMM matrix. Such a portrait can be misleading and difficult to interpret. For example, this practice ignores the fact that validity coefficients can be inflated by response bias and shared error variance. With regard to criterion-based, heterogeneous occupational scales, the method factor is increased by scoring the same item in several different scales, thus contributing correlated error variance. In the present study, inspection of the validity diagonals indicated that the average correlation ($r = .45$) between matched occupational scales on the CISS, KOIS, and SII for Personnel Managers ($r = .41$), Psychologists ($r = .54$), and Counselors ($r = .33$) was somewhat higher than those reported in four previous studies (O'Shea & Harrington, 1971; Triggs, 1943, 1944; Zytowski, 1972b) but similar to those reported by Carek (1972). Still, these matched scales differ enough to renew our concerns about their comparability. Again, the KOIS appears to be the outlier. When the KOIS correlations are removed, the median correlation between the three occupational scales on the SII and CISS was noticeably higher ($r = .56$). While the three occupational scales demonstrated the predicted convergent validity values, the results indicate that method variance strongly affects the KOIS. All three KOIS occupational scale method correlations substantially exceeded the convergent validity values. The correlation between the Personnel Manager and Psychologist scales was $r = .63$; between the Personnel Manager and Guidance Counselor scales, it was $r = .70$; and between the Psychologist and Guidance Counselor scales, it was $r = .74$. Method variance also appears to strongly affect the SII School Counselor scale given that it correlates nearly the same to the SII Human Resource Director scale ($r = .65$) and CISS Guidance Counselor scale ($r = .67$) and much lower to the KOIS Guidance Counselor scale ($r = .25$). Error variance did not appear to affect the scales on the three inventories. The MTMM results for the three matched scales indicate that the current versions of these inventories demonstrate the same general pattern of validity coefficients as reported previously reported. Furthermore, this study shows for the first time that discriminant validity evidence in the form of convergent validity exceeding control values is generally strong for the occupational scales but that method variance can be problematic in general for some interest inventories and in particular for some occupational scales on any inventory.

Vocational Self-Rating Scales

Table 8 shows the MTMM matrix that was constructed to assess the convergent and discriminant validity of the CISS Self-Efficacy scales, the SII Skills Confidence

TABLE 8
MTMM of Ability Self-Estimates on the CISS, SDS, and SII

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
CISS Skills																									
1. Adventuring																									
2. Producing	.44																								
3. Analyzing	.00	.23	—																						
4. Creating	.18	.35	.06	—																					
5. Helping	.13	.12	.07	.30																					
6. Influencing	.34	.28	.05	.55	.36																				
7. Organizing	.28	.36	.26	.30	.11	.53	—																		
SDS Competence																									
8. Realistic	**.23**	**.56**	.23	.07	.16	.02	.04																		
9. Investigative	.02	.06	**.68**	.04	.07	−.08	.08	.34	—																
10. Artistic	.12	.22	.09	**.50**	.17	.21	.06	.18	.21	—															
11. Social	.14	.21	.31	.15	**.40**	.05	.03	.33	.22	.01	—														
12. Enterprising	.25	.11	.08	.16	.01	**.56**	.36	.20	.08	.19	.20	—													
13. Conventional	.16	.19	.25	.09	.06	.14	**.54**	.16	.23	.04	.03	.28	—												
SDS Self-Estimates																									
14. Realistic	**.17**	**.45**	.02	.03	.27	.07	.08	**.58**	.01	.09	.21	.15	.10												
15. Investigative	−.02	.07	**.78**	.00	−.12	−.09	.11	.26	**.78**	.12	−.24	.09	.17	.06											
16. Artistic	.13	.31	.07	**.37**	.00	.03	−.02	.30	.11	**.68**	−.17	.04	−.12	.23	.15	—									
17. Social	.18	.02	.10	.13	**.33**	.00	.07	.04	.09	.27	**.42**	.22	.07	.06	.06	.23									
18. Enterprising	.24	.19	.033	.31	.08	**.57**	.47	.15	.04	.07	.14	**.63**	.18	.11	.03	.01	.16								
19. Conventional	.00	.06	.05	.01	.03	.14	**.58**	−.14	.00	.01	.14	.21	**.61**	.05	.06	−.13	.08	.27	—						
SII Skills Confidence																									
20. Realistic	**.11**	**.47**	.27	.08	.03	.07	.15	**.48**	.15	.09	.12	.14	.14	**.50**	.15	.09	.09	.02	.01						
21. Investigative	.04	.13	**.79**	.11	−.04	−.01	.14	.27	**.66**	.12	−.34	.05	.07	.09	**.77**	.11	.01	−.01	−.04	.45	—				
22. Artistic	.16	.35	.14	**.61**	.18	.32	.03	.34	.15	**.68**	−.19	.06	−.11	.20	.09	**.61**	.13	−.01	−.26	.40	.27	—			
23. Social	.18	.01	.02	.30	**.57**	.27	.04	.08	.12	.20	**.32**	.09	.17	.23	.01	.06	**.43**	.04	.00	.05	.10	.25	—		
24. Enterprising	.20	.32	.11	.54	.19	**.70**	.42	.33	.06	.34	.19	**.49**	.14	.24	.06	.30	.01	**.46**	.11	.21	.13	.40	.38	—	
25. Conventional	.00	.16	.51	.16	−.02	.17	**.58**	.16	.23	.10	−.17	.11	**.53**	.23	.33	.03	.03	.09	**.47**	.33	.54	.10	.20	.32	—

scales, the SDS Competence scales, and the SDS Ability Self-Estimate scales. Although the matrix of correlations among the four sets of self-ratings includes results for the CISS Adventuring scale, our analyses focused on the six RIASEC constructs and excluded the Adventuring scale. Thus, before considering the RIASEC blocks, we examined the Adventuring scale separately.

As with the vocational interest scales, the CISS Adventuring self-efficacy scale failed to converge with the SII Realistic self-efficacy scale, correlating a nonsignificant $r = .17$. The CISS Adventuring self-efficacy scale demonstrated higher method values than validity values in correlating higher than $r = .17$ to four of the six remaining CISS scales, namely $r = .44$ to Producing self-efficacy, $r = .34$ to Influencing self-efficacy, $r = .28$ to Organizing self-efficacy, and $r = .18$ to Creating self-efficacy. Relative to the SDS Competence scales, the Adventuring convergent validity value of $r = .23$ was exceeded by a correlation of $r = .25$ between the CISS Adventuring self-efficacy and SDS Enterprising competence scales. Relative to the SDS Ability Self-Estimate scales, the CISS Adventuring convergent validity value of $r = .17$ was exceeded by correlations of $r = .24$ to Enterprising and $r = .18$ to Social. The validity value of $r = .17$ between Adventuring and the SII Realistic skills confidence scales was exceeded by four of the six method correlations (Producing $r = .44$, Influencing $r = .34$, Organizing $r = .28$, and Creating $r = .18$).

CISS Skills

In examining the MTMM for the CISS Skills scales, we first considered convergent validity by inspecting parallel correlations among the CISS Skills, SII Skills Confidence, SDS Competence, and SDS Ability Self-Estimate scales. The 18 monotrait–heteromethod correlations were each statistically significant at the .01 level. They ranged from $r = .33$ between CISS Helping skill and SDS Social ability self-estimates to $r = .79$ between CISS Analyzing skill and SII Investigative skills confidence. The median correlation coefficient for the CISS Skills convergent validities to the SII Skills Confidence scales was $r = .60$ compared to $r = .55$ to the SDS Competence scales and $r = .51$ to the SDS Ability Self-Estimate scales.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the CISS validity values exceeded the control correlations, which ranged from $r = -.31$ between CISS Analyzing skill and SDS Social competence to $r = .54$ between CISS Creating skill and SII Enterprising skills confidence. Each of the validity values exceeded all of its control values. We should note that 20 of the 90 heterotrait–heteromethod correlations were significant. Nevertheless, compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 4.15$, $p < .001$). This shows good evidence of discriminant validity. The highest significant correlations occurred between CISS Creating skill and SII Enterprising skills confidence ($r = .54$) and between CISS Analyzing skill and SII Conventional skills confidence ($r = .51$). CISS Organizing skill correlated $r = .42$

to SII Enterprising skills confidence and $r = .42$ to SDS Enterprising ability self-estimates.

Relative to the second discriminant validity criterion, we determined whether the validity values exceeded method values, which for the CISS Skill scales ranged from $r = -.07$ between the Analyzing and Helping scales to $r = .55$ between the Creating and Influencing scales. With regard to the six RIASEC skill constructs measured by the CISS, one of the six scales had method correlations that exceeded its convergent validity correlations. The Helping and SDS Enterprising ability self-estimate scales correlated $r = .36$ compared to a validity value of $r = .33$ between CISS Helping and SDS Social ability self-estimate. The method correlations ranged from $r = -.07$ between the Analyzing and Helping scales to $r = .55$ between the Creating and Influencing scales. However, 10 of the 15 CISS heterotrait–monomethod correlations were significant. Compared to the monotrait–heteromethod correlations, proportionately there were no differences between the number of significant convergent and method correlations ($z = 1.62$, $p = .11$), suggesting that method variance may affect CISS Skill scores.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the CISS Skill scales by comparing the magnitude and pattern of correlations. With regard to magnitude, the median control correlation was $r = .15$ for the SII, $r = .16$ for the SDS Competence scales, and $r = .03$ for the SDS Ability Self-Estimate scales, compared to a median method correlation of $r = .28$ for the CISS. With regard to the pattern of correlations, the coefficient of concordance was $W = .64$ ($p < .001$). The concordance between just the CISS and SII was $W = .79$ ($p < .003$). Taken together, the relative magnitude and pattern similarity indicators suggest some response bias in the CISS Skill scales yet still meets the third criterion for discriminant validity.

SDS Competence

In examining the MTMM for the SDS Competence scales, we first considered convergent validity by inspecting the correlations between SDS Competence scales and parallel scales for CISS Skill, SII Skills Confidence, and SDS Ability Self-Estimates. The 18 monotrait–heteromethod correlations were each statistically significant at the .01 level. They ranged from $r = .32$ between SDS Social competence and SII Social skills confidence to $r = .79$ between SDS Investigative competence and SII Investigative skills confidence. The median correlation coefficient for the SDS Competence scales convergent validities to the SDS Ability Self-Estimate scales was $r = .62$ compared to $r = .55$ to CISS Skill and $r = .51$ to SII Skills Confidence.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the SDS Competence scale validity values exceeded its control values, which ranged from $r = -.34$ between SDS Social competence and SII Investigative skills confidence to $r = .34$ between SDS Artistic competence and SII Enterprising skills confidence as well as $r = .34$ between SDS Realistic competence and SII Artistic skills

confidence. Each of the validity values exceeded all of its control values, with a single exception. The SDS Social competence scale's correlation of $r = .32$ to the SII Social skills confidence scale was exceeded by four control correlations: SDS Enterprising competence to CISS Organizing skill ($r = .36$), SDS Artistic competence to SII Enterprising skills confidence ($r = .34$), SDS Realistic competence to SII Artistic skills confidence ($r = .34$), and SDS Realistic competence to SII Enterprising skills confidence ($r = .33$). We should note that 19 of the 90 heterotrait–heteromethod correlations were significant. Nevertheless, compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 4.26$, $p < .001$). This provides support for discriminant validity.

Relative to the second discriminant validity criterion, we determined whether the SDS Competence scale validity values exceeded its method values, which ranged from $r = -.33$ between the Realistic and Social scales to $r = .34$ between the Realistic and Investigative scales. One of the method correlations exceeded the convergent validity correlations. The SDS Social competence scale correlated $r = .32$ to SII Social skills confidence but was exceeded by a method correlation of $r = .34$ to both the SDS Realistic competence scale and the SDS Investigative competence scale. Of the 15 method correlations, 7 were significant for the SDS Competence scales. Nevertheless, compared to the monotrait–heteromethod correlations, proportionately there was a significant difference between the number of significant convergent and method correlations ($z = 2.27$, $p = .023$), suggesting that method variance did not adversely affect SDS Competence scale scores.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the SDS Competence scales by comparing the magnitude and pattern of correlations. With regard to magnitude, the median control correlation was $r = .07$ for CISS Skill, $r = .12$ for SII Skills Confidence, and $r = .09$ for SDS Ability Self-Estimates compared to a median SDS Competence scale method correlation of $r = .18$. Regarding the pattern of correlations, the coefficient of concordance was $W = .41$ ($p < .001$). Taken together, the relative magnitude and pattern similarity indicators suggest the presence of some response bias attributable to measurement methods, yet the SDS Competence scales still meet the third criterion for discriminant validity.

SDS Ability Self-Estimate Scales

In examining the MTMM for the SDS Ability Self-Estimate scales, we first considered convergent validity by inspecting parallel correlations between the SDS Ability Self-Estimate scales and the CISS Skill, SII Skills Confidence, and SDS Competence scales. The 18 monotrait–heteromethod correlations all were statistically significant at the .01 level. They ranged from $r = .33$ between SDS Social ability self-estimate and CISS Helping skill to $r = .78$ between SDS Investigative ability self-estimate and SDS Investigative competence. The median convergent validity coefficient for the SDS Ability Self-Estimate scales correlated to the SDS Competence scales was $r = .62$ compared to $r = .51$ to the CISS Skill scales and $r = .49$ to the SII Skills Confidence scales.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the SDS Ability Self-Estimate validity values exceeded the control values, which ranged from $r = -.27$ between the SDS Realistic ability self-estimate scale and CISS Helping skill to $r = .47$ between the SDS Enterprising ability self-estimate scale and CISS Organizing skill. Each of the validity values exceeded all of its control values, with a single exception. The SDS Enterprising ability self-estimate scale's $r = .46$ correlation to SII Enterprising skills confidence was exceeded by its $r = .47$ correlation to CISS Organizing skill. We should note that 14 of the 90 heterotrait–heteromethod correlations were significant. Nevertheless, compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 4.93$, $p < .001$). This shows evidence of discriminant validity.

Relative to the second discriminant validity criterion, we determined whether the SDS Ability Self-Estimate validity values exceeded method values, which ranged from $r = -.13$ between the Artistic and Conventional scales to $r = .27$ between the Enterprising and Conventional scales. None of the method correlations exceeded the convergent validity correlations. Of the 15 method correlations, 3 were significant for the SDS Ability Self-Estimate scales. Compared to the monotrait–heteromethod correlations, proportionately there was a significant difference between the number of significant convergent and method correlations ($z = 3.35$, $p < .001$), suggesting that method variance did not affect SDS Ability Self-Estimate scores.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the SDS Ability Self-Estimate scales by comparing the magnitude and pattern of correlations. With regard to magnitude, the median control correlation was $r = .03$ for CISS Skill, $r = .03$ for SII Skills Confidence, and $r = .09$ for SDS Competence, compared to a median SDS Ability Self-Estimate method correlation of $r = .06$. Regarding the pattern of correlations, the coefficient of concordance was $W = .46$ ($p < .001$). Taken together, the relative magnitude and pattern similarity indicate some minimal response bias, yet the SDS Ability Self-Estimate scales still meet the third criterion for discriminant validity.

SII Skills Confidence Scales

In examining the MTMM for the SII Skills Confidence scales, we first considered convergent validity by inspecting the parallel correlations between SII Skills Confidence and the CISS Skill, SDS Competence, and SDS Ability Self-Estimate scales. The 18 monotrait–heteromethod correlations all were statistically significant at the .01 level. They ranged from $r = .32$ between SII Helping skills confidence and SDS Social competence to $r = .79$ between CISS Analyzing skill and SII Investigative skills confidence. The median correlation coefficient of the SII Skills Confidence scale convergent validities to the CISS Skill scales was $r = .60$ compared to $r = .62$ to the SDS Competence scales and $r = .49$ to the SDS Ability Self-Estimate scales.

After examining the evidence for convergent validity, we considered in turn each of the three criteria for discriminant validity. First, we determined whether the SII validity values exceeded the control values, which ranged from $r = -.34$ between SII Investigative skills confidence and SDS Social competence to $r = .54$ between SII Enterprising skills confidence and CISS Creating skill. Each of the validity values exceeded all of its control values, with a single exception. The SII Enterprising skills confidence scale's $r = .54$ control correlation to the CISS Creating skill scale exceeded its validity correlation of $r = .49$ to the SDS Competence scale. We should note that 21 of the 90 heterotrait–heteromethod correlations were significant. The highest significant correlations were between SII Enterprising skills confidence and CISS Creating skill ($r = .54$), between SII Conventional skills confidence and CISS Analyzing skill ($r = .51$), and between SII Enterprising skills confidence and CISS Organizing skill ($r = .41$). Compared to the monotrait–heteromethod correlations (validity values), proportionately fewer significant heterotrait–heteromethod correlations (control values) were observed ($z = 4.04$, $p < .001$). This shows good evidence of discriminant validity for all but the Enterprising scale.

Relative to the second discriminant validity criterion, we determined whether the SII skills confidence validity values exceeded its method values, which ranged from $r = -.05$ between the Realistic and Social scales to $r = .54$ between the Investigative and Conventional scales. One of the method correlations exceeded its convergent validity correlation. The SII Social scale correlated $r = .32$ to the SDS Social competence scale but $r = .38$ to the SII Enterprising skills confidence scale. The other validity values for SII Social skills confidence were $r = .57$ to CISS Helping skill and $r = .43$ to SDS Social ability self-estimates. We should note that 11 of the 15 correlations were significant for the SII. Compared to the monotrait–heteromethod correlations, proportionately there was not a significant difference between the number of significant convergent and method correlations ($z = 1.41$, $p = .16$), suggesting that method variance may affect SII Skills Confidence scores.

Relative to the third criterion for discriminant validity, we looked for the presence of method variance in the SII Skills Confidence scales by comparing the magnitude and pattern of correlations. With regard to magnitude, the median control correlation was $r = .15$ for CISS Skill, $r = .09$ for SDS Ability Self-Estimates, and $r = .14$ for SDS Competence, compared to a median SII method correlation of $r = .27$. With regard to the pattern of correlations, the coefficient of concordance was $W = .52$ ($p < .001$). However, when compared to just the CISS Skill scales, there is a noticeable increase in concordance ($W = .77$, $p < .004$). Taken together, the relative magnitude and pattern similarity indicates the presence of some response bias attributable to measurement methods, yet the SII Skills Confidence scales still meet the third criterion for discriminant validity.

Discussion of Results for Vocational Self-Rating Scales

The self-ratings of RIASEC capabilities on the CISS, SDS, and SII showed a general pattern of convergent and discriminant validity similar to the pattern

found for vocational interests. The four sets of scales showed convergent and discriminant validity, but these validity values may have been inflated by method variance attributable to response set bias. Compared to the interest scales, there seem to be more problems apparent with the first generation of vocational self-rating scales in the CISS and SII and in the SDS Social competence scale.

The median convergent validity value for the four sets of self-rating scales was $r = .57$. We found only one research report with which to compare this finding. The median SDS Ability Self-Estimates correlated $r = .54$ to SII Skills Confidence compared to a correlation of $r = .53$ between SDS Ability Self-Estimates and an SII occupational skill scale in a study reported by Brown, Lent, and Gore (2000). Expectations about future performance indicated by the CISS Skill and SII Skills Confidence scales covaried more similarly, as did the self-evaluations indicated by the SDS Competence and Ability Self-Rating scales. The CISS and SII expectation scales shared about 40% of common variance, as did the two sets of SDS self-evaluation scales. The common variance between future efficacy expectations and current self-evaluations was roughly 30%. We conclude that this pattern reflects empirically the difference in temporal perspective between the constructs of vocational skill and ability self-ratings.

Brown et al. (2000) asserted that Ability Self-Estimates indicate normative judgments about past accomplishments and current abilities, whereas skill denotes expectations about future performance in specific contexts. This temporal distinction, along with the distinction that sociocognitive theory makes between general abilities and specific behaviors in particular situations, is implied by the subtle difference between the words "able" and "capable." Ability connotes having a quality that makes an action possible, whereas capability connotes having the quality needed for some specific activity. In other words, capabilities refer to specific tasks. Juang and Vondracek (2001) implied as much when they suggested that self-evaluations address judgments about possessing abilities, whereas expectations address beliefs about using capabilities to reach occupational goals. Vocational judgments and occupational expectations are linked in that expectations about future occupational performance are informed by judgments about current vocational abilities. Brown and his colleagues (2000) hypothesized that self-estimates influence interest formation through the skill beliefs to which they give rise. Despite the distinction between ability self-evaluations and capability expectations, it appears that for many practical purposes the SDS Ability Self-Estimate and Competence scales can function as general indicators of occupational capability expectations, whereas the CISS Skill scale and the SII Skills Confidence scale can function as general indicators of vocational ability evaluations.

Although not the focus of the current study, we note that the median correlation ($r = .55$) between vocational interests and vocational skill was roughly similar to the correlation of $r = .51$ reported by Donnay and Borgen (1999). This correlation between interests and skill seems relatively high given that the correlation among vocational interests scales was $r = .63$ in the current study. Of course, the strong relation between interests and skill does not preclude the possibility that

skill plays a unique role in explaining vocational aspirations and predicting occupational entry, as reported by Donnay and Borgen. Nevertheless, the possibility of considerable overlap between interests and skill, as noted by Tracey (1997), merits further investigation.

GENERAL DISCUSSION

Interest inventories use different methods to measure the same or similar constructs. This study systematically compared five interest inventories at four levels of comparison: (a) vocational interests measured by rationally based homogeneous scales; (b) occupational interests measured by criterion-based heterogeneous scales; and (c) scales that measure self-ratings of efficacy, ability, and skill for tasks associated with RIASEC themes. MTMM matrices were used to examine the convergent and discriminant validity of vocational interest, occupational interest, and vocational self-rating scales. The findings from the current study show somewhat stronger convergent validity across inventories than was the case in previous investigations. Of course, this generalization is limited by a small select sample restricted to career counseling practitioners and professors ($N = 118$). Results of this study suggest that these five interest inventories display a moderate degree of convergence. Vocational interests showed moderate median correlations. The occupational scales and self-estimate scales correlated somewhat weaker. Despite the sample size, we conclude that the occupational and vocational scales correlated empirically near the upper limits imposed by design differences and reliability constraints. Although reassuring from a psychometric perspective, the results do raise concerns about the comparability of similarly and same-named scales that are considered after discussing the particular conclusions that can be drawn from the results of this study.

Applying the Campbell and Fiske (1959) criteria to the results from the MTMM matrix for vocational interest indicated that the homogeneous scales on the five inventories demonstrated convergent and discriminant validity. Convergent validity requires that the same constructs, or constructs that share similar names, measured by the five inventories produce positive and substantial correlations. We used Holland's RIASEC typology to structure our examination of the MTMM matrices because it is the most popular and empirically supported framework for organizing vocational interests. We expected a median correlation of about $r = .50$ on matched RIASEC interests from different inventories based on the review of prior studies and because, in tapping a broader domain, RIASEC scales should have less error. Indeed, the median scale correlation among the five sets of six RIASEC scales in the MTMM was $r = .59$, a number suggesting at least moderate convergent validity across all five inventories. The estimated upper limit for these correlation coefficients, based on reliability estimates for each set of scales, ranged from $r = .68$ to $r = .77$. The correlations ranged from a low of $r = .08$ between the UNIACT-R Business Operations scale and the CISS Organizing scale to $r = .84$ between the CISS Analyzing scale and the SII Investigative scale. Correlations were higher between inventories with similar construction such as the SDS

summary scales and SII GOT scales that both were constructed or co-constructed by Holland. In general, findings revealed a moderate to high degree of convergent validity among the five interest inventories despite substantially different items, scaling, and norming. Although the scales correlated moderately, they were not so strongly related as to produce identical results.

We conclude that interest inventories that were designed to measure RIASEC types demonstrated the highest degree of convergence. The other interest inventories converged well and offered unique measures of Musical and Organizing interest. The CISS Organizing scale seems to be unique in measuring management interest in taking responsibility for arranging, organizing, and systematizing. Conventional interests concentrate not on organizing or managing a system but rather on maintaining systems by preparing and monitoring records, correspondence, and accounts. We hypothesize that office managers score higher on Organizing interests, while the clerical workers they supervise score higher on Conventional interests.

The Kuder remains the only major criterion-based inventory that employs a forced-choice format and does not employ a general reference sample (as do the CISS and SII). This approach offers useful information that is neither provided in the other inventories nor highlighted in the RIASEC model. For example, because he used forced-choice items, Kuder was able to make rather specific distinctions among art, music, and literature. When the format is not forced choice, these items are highly correlated and compose only one scale (Guilford, Christensen, Bond, & Sutton, 1954), possibly explaining why the other four inventories in this study do not have separate scales for these aesthetic activities. Thus, because of its design, the Kuder scales do make useful distinctions within the Artistic domain, and these distinctions can make the KOIS the inventory of choice for some clients. However, while the distinctions can be worth making, the apparent independence among these scales can be misleading. The same logic applies to the two Conventional scales that measure Clerical and Computational interests as well as to the two Realistic scales that measure Outdoor and Mechanical interests. Guilford and his colleagues (1954) suggested that if an inventory constructor wanted to measure distinctions within the Artistic domain, two factors that seem to actually be independent are artistic expression and artistic appreciation. To date, none of the major interest inventories has scales that distinguish appreciation of aesthetic products from interest in producing them.

Implications for Theory

The results of the current study have implications for both the science of vocational psychology and the practice of career counseling. First, let us consider implications for the psychology of interests. In a long-forgotten yet still important book titled *Psychological Research,* Underwood (1957) contrasted the literary explanation of a construct to its operational definition. Literary explanations assign conceptual meaning and thereby communicate and generate science. By contrast, operational definitions concretize meaning and thereby enable the conduct of

science. The current study again clearly showed the differences between operational designations of constructs with the same linguistic design. Each of the five interest inventories in the current study includes scales that purport to be operational indicators of RIASEC interests as set forth in linguistic form by Holland's theory of vocational personality types. The results of the current study indicate that, because the scales share only about 35% common variance, counselors cannot equate Holland's linguistic explanations with the operational definitions provided by the inventories. This means that, when using a particular inventory, counselors must be aware of how it operationally defines the specific constructs that it purports to measure, and they must remember that an individual scale on the inventory is just one variation on the theme of meaning explicated by the literary conception of the construct being measured. For example, counselors should not equate Realistic interests with the SDS Realistic scale, which is just one operational definition of the construct. The literary conception of Realistic interests benefits from multiple operational definitions that include mechanical, outdoor, and productive activities, while Conventional interests benefit from multiple operational definitions that separate organizing office systems from clerical routine. In this study, we found multiple operational definitions to be important for several constructs, but none more so than Artistic interests. Counselors should not assume that Artistic scales on interest inventories reflect response dispositions toward art, music, and literature. Only the KOIS seemed to attend specifically to musical interests, and the other inventories tended to emphasize either art or literary activities as manifestations of Artistic interests.

Now may be the time to explicate linguistically the facets of RIASEC interests, much as personality psychologists have explicated linguistically facets for each of the "Big Five" personality factors. For instance, Artistic interests may have facets involving the fine arts, functional crafts, writing, music, literature, aesthetics, acting, architecture, and philosophy. Of course, these facets are already implicit in Holland's theory of vocational personality types and may be explicit in his occupational classifications system in the form of secondary codes that follow "A." For example, "AI" may signify architecture, "AS" may signify writing, "AE" may signify music, "AC" may signify editing or collecting, and "AR" may signify crafts. A high peak on "A" alone may signify art. Clearly, counselors must be highly skilled in using secondary and tertiary codes to further specify the meaning of Artistic interests. Nevertheless, it could advance interest theory if we now use secondary codes to conceptualize linguistic explanations of facets for each RIASEC type. Work has already begun to conceptualize facets for occupational scales. Majors and Larson (2001) recommended identifying the underlying dimensions that are specific to individual occupational scales of the SII and did so with one scale: the female and male Business Education Teacher. They factor-analyzed the scale items, which are the SII items that differentiate business education teachers from the general reference samples. The resulting factors indicated the underlying dimensions that actually represent the multiple interests of business education teachers. Majors and Larson concluded that these underlying dimensions are for the most

part RIASEC constructs, yet a unique factor did emerge for both the male and female scales, one that seemed related to differences between male and females who teach business education courses.

Vocational psychologists have done a good job of linguistically conceptualizing interest facets for the RIASEC types and dimensions for criterion-based, heterogeneous occupational scales. They should now use the accumulated research on operational definitions of vocational and occupational interests to reexamine and refine the linguistic conceptions of the RIASEC constructs and explicitly link these literary conceptions to the operational definitions that best denote them. This activity would use the interaction between concepts and data to sharpen our ideas about vocational and occupational interests and about how to measure them.

Implications for Practice

While researchers and theorists advance the scientific understanding of interests, counselors should probably remain concerned about scales that measure 22 same- or similarly named occupational interests on the CISS, KOIS, and SII had a median correlation of $r = .42$ and that scales that measure RIASEC vocational interests on the CISS, KOIS, SDS, SII, and UNIACT-R had a median correlation of $r = .59$. Accordingly, counselors may wish to consider returning to a practice engaged in by the previous generations of career counselors. During the middle third of the 20th century, many counselors used more than one interest inventory with each client. They selected inventories that complemented each other in the domains measured and the scaling methods used. This practice resulted in multiple operational definitions of each client's vocational and occupational interests, pointing to a convergent validity for that client and improving the interpretive validity of each profile. It was common for a vocational appraisal battery to include both the Kuder Preference Record and the Strong Vocational Interest Blank. This practice started to fade as the costs of the inventories increased. Few centers can afford to use more than one inventory. Furthermore, the test constructors made it more attractive to use a single inventory by adding different types of scales to existing measures. For example, the SVIB began including Basic Interest Scales, the Strong–Campbell Interest Inventory started to report RIASEC scores, and the KOIS reported both vocational interest estimates and occupational scores. However, as the results of this study indicate, the possibility exists that using one interest inventory might only partially assess a client's interests and may even ignore measuring particular interests (e.g., musical interests, organizing interests) that could be critical in the client's decision making. If nothing else, the results suggest that counselors should consider having available several interest inventories from which to select the most appropriate one to administer to each client rather than routinely giving the same inventory to every client (Spokane, 1991). Even when administering inventories to large groups, school counselors might consider administering a vocational interest inventory to 9th- and 10th-graders while administering an occupational interest inventory to 11th- and 12th-graders.

Recommendations for Research

Of course, the other point of view is to use the same inventory with every client so as to develop clinical skills for interpreting that inventory. The current study examined the convergent validity of the scales across inventories; we did not compare the similarity of the interpretations that the profiles generate. The validity of an interpretation cannot be established by the construct validity data reported herein. A study of interpretive validity would concentrate on how confident counselors can be in using inventories to measure interests, not on the convergent validity of inventory scales or using interest inventories to predict educational/vocational choices and occupational entry. In many respects, interpretive validity is a more crucial question when decisions must be made about comparability across inventories. Scales from different inventories may correlate only moderately yet still lead to similar conclusions when the results are interpreted by a counselor (i.e., interpretive validity). Indeed, our analysis of the profiles for Ellenore Flood, who completed the same five inventories examined in the current study, revealed high levels of both profile validity in the form of congruent summary codes and interpretive validity in the similarity of interpretations and counseling strategies employed by the counselors who wrote about them (Savickas, 1998; Spokane, 1998). However, we need to study profiles from many clients to thoroughly examine the issues of profile validity and interpretive validity.

Studies of profile validity can start by investigating whether different inventories produce similar vocational interest profiles. This would involve taking the person approach to research (Reitzle & Vondracek, 2000) and determining the similarity of RIASEC summary codes produced by each of the five inventories for each person. Kuder (1969) suggested that a way to study profile validity for occupational scales would be to determine the efficiency with which two inventories distinguish between a pair of occupations, as indicated by the proportion of workers that each inventory correctly classifies. Studies of interpretive validity would involve having a group of counselors interpret the results from sets of five inventories for those individuals who show the most dissimilar results. Such studies should focus on examining the extent to which two interest inventories lead to similar results in a counseling situation (Kuder, 1969). As Tinsley and Chu (1999) concluded in their incisive review of test interpretation outcomes, knowledge in this area is "shockingly inadequate." Fortunately, the clinical use of interest inventories is an important area that must be examined directly given the recent focus on counseling process studies that investigate the clinical use of interest inventories to "diagnose" interests that further clients' adaptation as opposed to the use of inventories to "create" interests that foster clients' development (Spokane, 1998).

Limitations

A clear limitation of the current study is the small and homogeneous sample used. The participants in this study consisted of career counseling practitioners and professors. This could have led to a restriction of range in the scores and their resulting correlations. A larger and more diverse sample might yield somewhat different

results. Nevertheless, we believe that this homogeneous group of participants provided a study that controlled for variable occupations. The restriction in range was consistent across all variables, and this causes less difficulties in interpreting the results than would different degrees of restriction for different variables (Campbell & Fiske, 1959, p. 102). A delimitation of the study was that it concentrated on the adequacy of interest inventories as measures of vocational and occupational interests; it did not concentrate on the adequacy of the construct of interests by investigating correlations with other constructs. The literature on the concurrent and predictive validity of these inventories already shows their effectiveness in facilitating and predicting vocational choice and occupational entry.

Summary

The results of the study add evidence to the argument that interest inventories provide only a partial indicator of vocational interests. The inventories seem to work differently but probably equally well. Those inventories constructed using similar methods, such as the CISS and SII as well as the SDS and UNIACT-R, correlate higher. Inventories designed to measure RIASEC types have vocational interest scales that correlate higher because they seek to operationally define the same linguistic concepts. Homogeneous scales that measure vocational interests correlate higher than do heterogeneous scales that measure the similarity of interests to those of members of occupational groups. Among occupational interest scales, those with more homogeneous groups of incumbents correlate higher and may work better. Different inventories operationally define different literary conceptions of psychological constructs such as Realistic and Artistic interests. Indeed, no "gold standard" exists for vocational interest measurement. At this juncture in our development, the diversity of operational definitions, including a variety of scaling and norming techniques, is valuable for advancing science and diversifying practice. As vocational psychology struggles to redefine its literary conception of vocational interests (Savickas, 1999), the inconsistency in different operational definitions, reaffirmed by the results of this study, can be considered useful variance in prompting insights into the meaning of vocational and occupational interests and the linguistic explanation of interest types.

REFERENCES

American College Testing. (1995). *Technical manual: Revised unisex edition of the ACT Interest Inventory (UNIACT)*. Iowa City, IA: ACT Inc.

Best, S., & Knapp-Lee, L. (1982). Relationship of interest measurement derived from the COPS System Interest Inventory and the Kuder DD General Interest Survey: Construct validation of two measures of occupational activity preferences. *Educational and Psychological Measurement*, **42**, 1289–1293.

Betz, N. E., Borgen, F. H., & Harmon, L. W. (1994). *Skills Confidence Inventory: Applications and technical guide*. Palo Alto, CA: Consulting Psychologists.

Betz, N. E., Borgen, F. H., Kaplan, A., & Harmon, L. W. (1998). Gender and Holland type as moderators of the validity and interpretive utility of the Skills Confidence Inventory. *Journal of Vocational Behavior*, **53**, 281–299.

Brown, S. D., Lent, R. W., & Gore, P. A., Jr. (2000). Self-rated abilities and self-efficacy beliefs: Are they empirically distinct? *Journal of Career Assessment,* **8,** 223–235.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin,* **56,** 81–105.

Campbell, D. P., Hyne, S. A., & Nilsen, D. L. (1992). *Manual for the Campbell Interest and Skills Inventory.* Minneapolis, MN: National Computer Systems.

Carek, R., (1972). Another look at the relationship between similar scales on the Strong Vocational Interest Blank and Kuder Occupational Interest Survey. *Journal of Counseling Psychology,* **19,** 218–223.

Clark, K. E. (1961). *Vocational interests of nonprofessional men.* Minneapolis: Univ. of Minnesota Press.

Cronbach, L. J. (1980). Validity on parole: How can we go straight? *New Directions for Testing and Measurement,* **5,** 99–108.

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper & Row.

Diamond, E. E. (1990). The Kuder Occupational Interest Survey. In C. E. Watkins, Jr., & V. L. Campbell (Eds.), *Testing in counseling practice* (pp. 211–239). Hillsdale NJ: Erlbaum.

Dolliver, R. H., & Nelson, R. E. (1975). Assumptions regarding vocational counseling. *Vocational Guidance Quarterly,* **24,** 12–19.

Donnay, D. A. C., & Borgen, F. H. (1999). The incremental validity of vocational self-efficacy: An examination of interests, self-efficacy, and occupation. *Journal of Counseling Psychology,* **46,** 432–447.

Guilford, J. P., Christensen, P. R., Bond, N. A., Jr., & Sutton, M. A. (1954). A factor analysis of human interests. *Psychological Monographs: General and Applied,* **68**(4), Whole No. 375.

Hansen, J. I. C., & Neuman, J. L. (1999). Evidence of concurrent prediction of the Campbell Interest and Skills Survey (CISS) for major college selection. *Journal of Career Assessment,* **7,** 239–247.

Harmon, L. W. (1999). Measuring interests: Approaches and issues. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 171–191). Palo Alto, CA: Davies–Black.

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

Holland, J. L. (1977). *Vocational Preference Inventory.* Palo Alto, CA: Consulting Psychologists.

Holland, J. L., Fritzsche, B. A., & Powell, A. B. (1994). *Technical manual for the Self-Directed Search.* Odessa, FL: Psychological Assessment Resources.

Juang, L., & Vondracek, F. W. (2001). Developmental patterns of adolescent capability beliefs: A person approach. *Journal of Vocational Behavior,* **59,** 34–52.

King, P., Norrel, G., & Powers, G. P. (1963). Relationships between twin scales on the SVIB and the Kuder DD. *Journal of Counseling Psychology,* **10,** 395–410.

Knapp, L., & Knapp, R. R. (1985). *California Occupational Preference System Interest Inventory.* San Diego: EdITS.

Knowles, E. R., & Condon, C. A. (2000). Does the rose still smell as sweet? Item variability across test forms and revisions. *Psychological Assessment,* **12,** 245–252.

Kuder, F. (1969). A note on the comparability of occupational scores from different interest inventories. *Measurement and Evaluation in Guidance,* **2,** 94–100.

Kuder, F. (1977). *Activity interests and occupational choice.* Chicago: Science Research Associates.

Kuder, F., & Zytowski, D. G. (1991). *Kuder Occupational Interest Survey Form DD general manual.* Adel, IA: National Career Assessment Associates.

Majors, M. S., & Larson, L. M. (2001). Occupationally unique interests in scales in the Strong Interest Inventory. *Journal of Vocational Behavior,* **59,** 105–119.

Maxwell, A. E. (1961). *Analysing qualitative data.* London: Methuen.

Omizo, M. M., & Michael, W. B. (1983). Relationship of COPS System Interest Inventory scales to Vocational Preference Inventory (VPI) scales in a college sample. *Educational and Psychological Measurement,* **43,** 595–601.

O'Shea, A. J., & Harrington, T. F., (1971). Using the Strong Vocational Interest Blank and the Kuder Occupational Interest Survey, Form DD, with the same clients. *Journal of Counseling Psychology,* **18,** 44–50.

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior,* **25,** 304–315.

Reitzle, M., & Vondracek, F. W. (2000). Methodological avenues for the study of career pathways. *Journal of Vocational Behavior,* **57,** 445–467.

Savickas, M. L. (1998). Interpreting interest inventories: A case example. *Career Development Quarterly,* **42,** 307–310.

Savickas, M. L. (1999). The psychology of interests. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 19–56). Palo Alto, CA: Davies–Black.

Spokane, A. R. (1991). *Career intervention.* Englewood Cliffs, NJ: Prentice Hall.

Spokane, A. R. (1998). Risk versus reluctance: Understanding an ambivalent entrepreneur. *Career Development Quarterly,* **46,** 370–375.

Spokane, A. R., & Holland, J. L. (1995). The Self-Directed Search: A family of self-guided career interventions. *Journal of Career Assessment,* **3,** 373–390.

Spokane, A. R., & Jacob, E. J. (1996). Career and vocational assessment 1993–1994: A biennial review. *Journal of Career Assessment,* **4,** 1–32.

Strong, E. K., Jr. (1943). *Vocational interests of men and women.* Stanford, CA: Stanford Univ. Press.

Tinsley, H. E. A., & Chu, S. (1999). Research on test and interest inventory outcomes. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 257–276). Palo Alto, CA: Davies–Black.

Tracey, T. J. G. (1997). The structure of interests and self-efficacy expectations: An expanded examination of the spherical model of interests. *Journal of Counseling Psychology,* **44,** 32–43.

Triggs, F. O. (1943). A study of the relation of Kuder Preference Record scores to various other measures. *Educational and Psychological Measurement,* **3,** 341–354.

Triggs, F. O. (1944). A further comparison of interest measurement by the Kuder DD Preference Record and the Strong Vocational Interest Blank for men. *Journal of Educational Research,* **37,** 538–544.

Wilson, R. N., & Kaiser, H. E. (1968). A comparison of similar scales on the SVIB and the Kuder DD, Form DD. *Journal of Counseling Psychology,* **15,** 468–470.

Zytowski, D. G. (1968). Relationships of equivalent scales on three interest inventories. *Personnel and Guidance Journal,* **47,** 44–49.

Zytowski, D. G. (1972a). Equivalence of the Kuder DD Occupational Interest Survey and the Strong Vocational Interest Blank revisited. *Journal of Applied Psychology,* **56,** 184–185.

Zytowski, D. G. (1972b). A concurrent test of accuracy-of-classification for the Strong Vocational Interest Blank and Kuder DD Occupational Interest Survey. *Journal of Vocational Behavior,* **2,** 245–250.

Zytowski, D. G. (1976). Predictive validity of the Kuder Occupational Interest Survey: A 12 to 19-year follow up. *Journal of Counseling Psychology,* **23,** 221–233.

Zytowski, D. G., & Hay, R. (1984). Do birds of a feather flock together? A test of the similarities within and the differences between five occupations. *Journal of Vocational Behavior,* **24,** 242–248.

Zytowski, D. G., & Laing, J. (1978). Validity of other gender normed scale on the Kuder Occupational Interest Survey. *Journal of Counseling Psychology,* **25,** 205–209.

Received September 7, 1999; published online May 16, 2002