Cognitive and Noncognitive Determinants and Consequences of Complex Skill Acquisition

Phillip L. Ackerman, Ruth Kanfer, and Maynard Goff
University of Minnesota, Twin Cities

Integration of multiple perspectives on the determinants of individual differences in skill acquisition is provided by examination of a wide array of predictors: ability (spatial, verbal, mathematical, and perceptual speed), personality (neuroticism, extroversion, openness, conscientiousness, and agreeableness), vocational interests (realistic and investigative), self-estimates of ability, self-concept, motivational skills, and task-specific self-efficacy. Ninety-three trainees were studied over the course of 15 hr (across 2 weeks) of skill acquisition practice on a complex, air traffic controller simulation task (Terminal Radar Approach Controller; TRACON; Wesson International; Austin, TX). Across task practice, measures of self-efficacy, and negative and positive motivational thought occurrence were collected to examine prediction of later performance and communality with pretask measures. Results demonstrate independent and interactive influences of ability tests and self-report measures in predicting training task performance. Implications for the selection process are discussed in terms of communalities observed in the predictor space.

In the past, studies aimed at prediction of performance over skill acquisition have typically examined only ability measures, only personality measures, or abilities and a few personality measures (e.g., for a review, see Kanfer, Ackerman, Murtha, & Goff, 1995). Moreover, no studies exist that jointly examine a wide array of cognitive ability (objective and perceived), affective (personality), and conative (motivational and interest) variables to estimate their independent contributions to prediction of individual differences in performance or the communalities among the constructs. The first purpose of the current investigation is to examine five domains of variables (personality, interests, self-concept, self-estimates of ability, and objective ability), as they independently and jointly predict performance over skill acquisition, along with variables of self-efficacy and motivation. The second purpose of this investigation is to examine the communalities among these domains of predictor variables and establish the grounds for an integrated perspective for the role that traits and dispositions have on individual differences in performance.

The approach adopted in this investigation is that the process of skill acquisition is dynamic—as performance changes over task practice, cognitive and noncognitive constructs may differentially determine individual differences in performance (e.g., see Ackerman, 1988, 1990, 1992; Kanfer & Ackerman, 1989). In addition, the task-relevant thoughts and self-efficacy expectations of trainees may...
change over the course of skill acquisition. We also examined how these changes occur, how trait measures predict individual differences in thoughts and self-efficacy over practice, and, most importantly, whether these thoughts and expectations, in turn, predict individual differences in subsequent task performance.

First, we briefly describe theories and data regarding the prediction of individual differences in performance during skill acquisition. Next, we describe theories and data regarding the effects of task practice on subsequent thoughts and self-efficacy expectations.

Prediction of Individual Differences in Performance

Abilities

Since the 1950s, substantial research attention has been devoted to documenting the changes in the ability determinants of task performance over practice (for reviews, see Ackerman, 1987; Adams, 1987; Fleishman, 1972). Since the 1970s, there have been several attempts to develop theories that predict the relative importance of different cognitive and intellectual abilities over the course of task practice (e.g., Ackerman, 1988; Kyllonen & Christal, 1989; Kyllonen & Woltz, 1989; for a review, see also Cronbach & Snow, 1977).

Two theories in particular relate to the current issues of task information-processing demands and the types of abilities that determine performance at different stages of practice, namely those of Ackerman (1988) and Kyllonen and Christal (1989). The model proposed by Kyllonen and Christal (1989) specifies four sources of individual differences in the acquisition of knowledge and procedural skills: breadth of declarative knowledge, breadth of procedural skills, capacity of working memory, and speed of processing. When a task is novel, the critical components of performance over task practice are working memory and processing speed, with working memory capacity most associated with initial performance and processing speed most associated with performance after practice (Woltz, 1988).

Ackerman's (1988) theory segments the ability—performance relations into three broadly defined stages of practice; namely, Fitts and Posner's (1967) cognitive, associative, and autonomous stages. The theory specifies that general and broad content abilities (spatial, verbal, and numerical) are most associated with novel task performance (cognitive stage), perceptual speed abilities are most associated with intermediate practiced performance (associative stage), and psychomotor abilities are most associated with performance after protracted practice (autonomous stage). The theory also specifies that tasks with only consistent information-processing components will show these three-stage changes in ability—performance relations (consistent refers to unchanging rules for mapping stimuli to responses). Inconsistent information-processing task components or tasks where the consistency of information processing is not apparent to the subjects will fail to show changes beyond the cognitive stage of practice (Ackerman, 1988; inconsistent refers to tasks where changes occur in mapping stimuli to responses). In addition, insufficient motivation to learn will also have the effect of precluding or slowing development of later stages of skill (Kanfer & Ackerman, 1989).

The theories make similar predictions for ability—performance relations during practice in several types of tasks because there is substantial overlap between measures of working memory capacity and measures of reasoning ability (Kyllonen & Christal, 1989). That is, initial task performance is well-predicted by general ability measures, which also include broad content tests (e.g., verbal, spatial, and numerical reasoning) and measures of working memory capacity (e.g., ABCD Order Test, Alphabet Recoding Test; see Woltz, 1988, for descriptions). At an intermediate stage of task practice, tests of perceptual speed ability (e.g., tests of substitution, clerical checking, and perceptual scanning) and tests of processing speed (e.g., encoding speed, retrieval speed, and response speed) are increasingly associated with individual differences in task performance (Ackerman, 1988; Woltz, 1988). The theories diverge in predicting the determinants of individual differences in task performance at the autonomous stage of skill acquisition. However, for the task under current consideration, a Terminal Radar Approach Controller (TRACON) simulation, and the amount of

1 TRACON is licensed software by Robert B. Wesson, Wesson International, Austin, Texas.
practice accorded (less than 20 hr) both theories make similar predictions of the importance of broad content abilities and perceptual speed abilities.

**Personality**

The history of personality test–performance relations in applied psychology lies beyond the scope of this article (but see, e.g., Kanfer et al., 1995). In an early review, Ellis and Conrad (1948, p. 421) concluded, “[P]ersonality inventories proved generally ineffective for predicting performance measures (such as successful completion of a training course).” However, recent studies have supported the conjecture that there are some replicable nontrivial correlations between personality measures and job behaviors. For example, studies by Borman, Rosse, and Abrahams (1980) and research conducted under the auspices of the U.S. Army Project A Team (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990) are representative of this approach. The Army Project A studies involved noncognitive and ability testing of over 10,000 personnel. In the Hough et al. (1990) report, analysis of the criterion space was conducted concurrently with investigation of the predictor domain to yield a model of personality–work criterion linkages. A noncognitive battery, designed to assess six predictor constructs, was developed and empirically evaluated in the context of relationships to a multidimensional set of performance criteria. McHenry et al. (1990) examined the predictive validity of the noncognitive measures for training and proficiency criteria. Results obtained in both studies indicate that two personality constructs, Dependability and Achievement, were found to be valid predictors of ancillary job performance criteria (e.g., Effort and Leadership and Personal Discipline).

For direct measures of performance, a related approach takes advantage of the Big Five factor taxonomy of personality constructs (that is, the five broad personality factors of Neuroticism, Extroversion, Openness, Agreeableness, and Conscientiousness; see e.g., Digman, 1990, or Goldberg, 1990). For example, Barrick and Mount (1991) examined the relationship of the Big Five personality dimensions to three job performance criteria by five occupational groups. Using applied studies appearing from 1952–1988, Barrick and Mount had raters classify the personality scales used in the research into the various predictor dimensions. Meta-analytic results indicate that only Conscientiousness was substantially related to performance across all occupations and for all types of criterion. In contrast, Extroversion was found to be a valid predictor across all criterion types for two occupations: management and sales.

**Motivational Skills**

Motivational theory (e.g., Kanfer & Ackerman, 1989, 1990, 1995; Kuhl, 1985) and previous empirical research with the selection of air traffic controllers (Ackerman & Kanfer, 1993a, 1993b) have suggested that level of two aspects of self-regulatory skills (namely, emotion control and motivation control) are important determinants of individual differences in performance over skill acquisition. In addition, a measure of these motivational skills has been shown to be independent of ability and provides incremental predictive validity for laboratory and field complex skill training.

**Self-Concept**

Modern conceptualizations of academic self-concept consider the construct to be competency-based and multifaceted with varying levels of specificity (Goff, 1994; Marsh, 1990). Such conceptualizations use judgments of competencies for performance of particular behaviors similar to self-efficacy judgments as their basis at the item level. Aggregation of items at differing levels of generality of self-concept is then possible. Several recent studies and reviews (Byrne, 1984; Marsh, 1990; Marsh, Byrne, & Shavelson, 1988; Norwich, 1987) have established that competency-based self-concept constructs have significant relations with matching areas of achievement in the academic domain. Although early reports on self-estimates of ability (e.g., DeNisi & Shaw, 1977; Kelso, Holland, & Gottfredson, 1977; Levine, Flory, & Ash, 1977) indicated limited communality with objective measures of ability, recent research has shown that self-concept measures (based on abili-
ties and competencies) and self-estimates of ability have convergent and discriminant associations with other constructs of interest in performance prediction such as ability, personality, and vocational interests (Goff, 1994).

**Vocational Interests**

Vocational interests can be thought of as a motivational representation of vocational self-concepts (Lowman, 1991). Several vocational interest and choice theories suggest that vocational interests involve self-concepts of competencies and motivational factors (Holland, 1973; Krumboltz, 1979; Krumboltz, Mitchell, & Jones, 1976, 1978; Super, 1957, 1963). Vocational interests have been shown to have differential associations with ability measures (for a review, see Ackerman & Goff, 1995). Such associations are thought to be based on the ability demands of job tasks clustered within each vocational interest area (Prediger, 1991), including abilities not typically tapped by ability tests. For instance, realistic job tasks are expected to have particular demands on spatial, mechanical, manual dexterity, and organization abilities. Specific task interest has been shown to predict learning performance (see Schiefele, 1991, for a discussion of the effects of content-specific task interest on learning performance). The combined representation of self-concepts of competencies and motivation within vocational interest assessments may result in useful prediction of learning of complex tasks that are similar to jobs.

**Self-Efficacy**

Self-efficacy judgments pertain to an individual's confidence in the ability to successfully mobilize and organize resources for goal attainment (Bandura, 1986). Conceptualized as task-specific and prospective judgments, self-efficacy may be only weakly related to individual differences in personality or other trait measures. Major determinants of integrative self-efficacy include (past) performance accomplishments, vicarious experiences, verbal persuasion, and physiological state. Numerous studies have provided support for the influence of these factors on self-efficacy judgments in achievement and clinical contexts (e.g., see Bandura, 1986).

According to social-cognitive theory, self-efficacy judgments may exert direct and indirect influences on subsequent learning and performance (see Bandura, 1986). Judgments of low self-efficacy are associated with lower levels of effort and lack of persistence; high self-efficacy is associated with higher levels of effort and persistence in the face of failure. Many studies show a positive association between self-efficacy judgments and subsequent task performance (e.g., for reviews see Bandura, 1986; Feltz, 1994; Sadri & Robertson, 1993). However, fewer studies have examined the unique influence of self-efficacy judgments on performance relative to other factors, such as cognitive abilities. Among studies reporting on investigation of the influence of self-efficacy judgments on performance relative to individual differences in cognitive abilities, interests, and other nonability traits, the pattern of results obtained is inconsistent. For example, Bandura and Wood (1989) and Mathieu, Martineau, and Tannenbaum (1993) provided evidence for the independent role of self-efficacy judgments on performance relative to the influence of initial task performance. However, Mitchell, Hopper, Daniels, George-Falvy, and James (1994) found no significant incremental influence of self-efficacy in the prediction of performance in a short skill acquisition experiment.

**Consequences of Skill Acquisition**

Previous research has indicated that there are numerous cognitive and affective consequences of skill acquisition (e.g., for a review see Kanfer & Ackerman, 1995). Generally, task self-efficacy expectations increase as skills are acquired, and problems associated with emotion control diminish, which would be reflected by a decrease in negative motivational thought frequency (Kanfer & Ackerman, 1989, 1990). In contrast, as skills are acquired, greater demands on motivation control are expected. Positive motivational thought frequencies may be expected to change in concert with positive self-regulatory processes. Whether these positive cognitions increase in frequency or remain stable is as yet unclear, given competing desires on the part of trainees to minimize effort expenditures and maximize performance levels (Kanfer, 1987).
Summary and Overview

Six different sources of individual differences determinants are considered: ability, personality, motivation, self-concept, interests, and self-efficacy. Although these sources may each offer significant prediction of complex skill acquisition, they are typically evaluated in isolation. The purpose of the experiment described here was to examine how these six sources of influence relate to individual differences in complex task acquisition both in isolation and in conjunction with one another. By jointly examining the various determinants of performance, we can evaluate the evidence for the relative independence or communality of the constructs for the purposes of efficiently predicting performance and understanding trait overlap.

Layout of Experimental Design and Overview of Measures

This experiment contained four major categories of measures: distal individual differences, proximal individual differences, concomitant proximal, and criterion (TRACON) task performance. These measures, described briefly below, are discussed in greater detail in the Method section.

1. Distal. Distal refers to measures that are general and are typically thought to represent trait constructs. Numerous distal measures were administered in order to converge on the multiple determinants of performance in the TRACON simulation task. They were designed to assess individual differences in cognitive-intellectual abilities, personality, vocational interests, and self-ratings of ability and self-concept (along with a short measure of motivational skills). In addition, the Dial Reading Test and a companion test, Directional Headings Test were administered to evaluate whether these measures had assessed common variance among ability and nonability influences on performance.

2. Proximal. Proximal refers to measures that are task-specific and are associated with a particular situational context. A questionnaire (Interim Questionnaire) was administered 1 or 2 days after trainees had viewed the instructions for performing TRACON and just prior to actual task engagement. This questionnaire assessed negative motivation and positive motivation thoughts directly pertaining to the TRACON task and several aspects of task-specific self-efficacy. This questionnaire was also administered prior to each subsequent TRACON session.

3. Concomitant-Proximal. Two types of concomitant measures were used during TRACON practice: an Interim Questionnaire was administered just prior to the beginning of each TRACON session and a Task Perceptions Questionnaire was administered just after each TRACON session. Both measures assessed task-specific retrospective and prospective thoughts. Retrospective thought frequency for the Interim Questionnaire pertained to intrusive (negative motivation) and purposeful (positive motivation) thought frequency during the days between TRACON sessions. For the Task Perceptions Questionnaire, retrospective thoughts pertained to the six simulation trials in the just completed session. Prospective thoughts pertained to TRACON task self-efficacy (both measures) and performance expectancies.

4. Task Performance. As in previous studies (Ackerman, 1992; Ackerman & Kanfer 1993b), the major measures used to assess performance in TRACON related to the number of planes successfully handled (carried to final disposition) during each simulation. Overall performance was measured, as were separate performance components of handling arrival flights and overflights.

Figure 1 shows the order of administration of the various measures during the study.

Method

Participants

Trainees were recruited by signup sheets and flyers distributed around the campus of the University of Minnesota. Trainees were restricted to individuals with (a) ages between 18 and 30, (b) normal or corrected-to-normal vision, hearing, and motor coordination, and (c) no prior experience with air traffic controller tasks. Ninety-three participated in the experiment. Of these trainees, 42 were men (age level for the whole sample was $M = 20.4$ and $SD = 2.75$). Trainees were paid $125 each for the 24 hr of participation, and this payment was not contingent on task or test performance.
COGNITIVE AND NONCOGNITIVE DETERMINANTS

Session 1 - Distal Measures (Ability)

<table>
<thead>
<tr>
<th>108 min</th>
<th>Aptitude Assessment Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Motivational Skills (self-report)</td>
</tr>
<tr>
<td>5</td>
<td>Break</td>
</tr>
<tr>
<td>53</td>
<td>Additional Ability Tests</td>
</tr>
</tbody>
</table>

Session 2 - Distal Measures (Self-Report)

<table>
<thead>
<tr>
<th>25 min</th>
<th>Personality</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Self-Estimates of Ability</td>
</tr>
<tr>
<td>6</td>
<td>Typical Intellectual Engagement</td>
</tr>
<tr>
<td>5</td>
<td>Break</td>
</tr>
<tr>
<td>7</td>
<td>Vocational Interests</td>
</tr>
<tr>
<td>3</td>
<td>Self-Concept</td>
</tr>
<tr>
<td>5</td>
<td>Break</td>
</tr>
<tr>
<td>85</td>
<td>TRACON Video Training Tape (Embedded Quizzes)</td>
</tr>
</tbody>
</table>

Sessions 3 - 8 (Proximal Measures and Task Practice)

<table>
<thead>
<tr>
<th>11 min</th>
<th>Interim Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>TRACON (two trials)</td>
</tr>
<tr>
<td>5</td>
<td>Break</td>
</tr>
<tr>
<td>60</td>
<td>TRACON (two trials)</td>
</tr>
<tr>
<td>5</td>
<td>Break</td>
</tr>
<tr>
<td>30</td>
<td>TRACON (one trial)</td>
</tr>
<tr>
<td>6</td>
<td>Task Perceptions Questionnaire</td>
</tr>
</tbody>
</table>

Figure 1. Order of administration of ability, self-report, and task measures during the study. TRACON = Terminal Radar Approach Controller Task.

Apparatus

The computerized self-report measures and the TRACON trials were administered on IBM or Compaq 80386 computers. Information was displayed in text mode (for the questionnaires) or in color VGA (visual graphics array; 680 horizontal × 480 vertical pixels) resolution graphics (for the TRACON trials). Input was accomplished by keyboard responses or with a MicroTech three-button trackball. Audio information (from TRACON) was presented binaurally over headphones to each trainee (with a SoundBlaster interface). Each computer was on a separate table and was shielded visually and aurally by large acoustic panels that enclosed each computer. For ability testing, trainees were divided into groups of up to 40. For TRACON task practice, trainees were divided into groups up to 20.

Ability Tests

Two batteries of tests were administered. The first was the Aptitude Assessment Battery (AAB), the development of which was described extensively in Ackerman and Kanfer (1993b). The second battery was used to supplement the AAB and to assess specifically perceptual speed and verbal abilities. The AAB contains nine ability tests and an 18-item measure of motivational skills in the following order: Necessary Facts, Spatial Orientation, Math Knowledge, Spatial Analogy, Problem Solving, Paper Folding, Verbal Test of Spatial Ability, Dial Reading Test, and Directional Headings. (The Motivational Skills measure follows the ability tests in the administration order). The supplemental battery contained six tests in the following order: Vocabulary, Letter/Number Substitution, Controlled Associations, Subtraction and Multiplication, Word Beginnings, and CA-2 (Clerical Ability). Descriptions of the AAB tests can be found in Ackerman and Kanfer (1993b). The three verbal tests and the Subtraction and Multiplication tests in the supplemental battery are from the Educational Testing Service Kit (Ekstrom, French, Harman, & Dermen, 1976). The Letter/Number Substitution test is a locally developed perceptual speed test (Ackerman, 1986), and the CA-2 is a clerical checking test (Bennett & Gelink, 1951). Together, these batteries were designed to assess the following four factors (Spatial Ability, Verbal Ability, Perceptual Speed [Simple], and Math ability). Composites for each ability are computed with unit-weighted z scores of the component tests, and an overall general ability composite is computed across all five factors.

In addition to the four standard ability factors, we also collected data on the Dial Reading Test (based on a description in Guilford & Lacey, 1947; see Ackerman & Kanfer, 1993b) and a similar measure, the Directional Headings Test. The Directional Headings Test is a test of memory, perceptual encoding, and learning that was developed by the Federal Aviation Administration Civil Aeromedical Institute (see Cobb & Mathews, 1972). Trainees are given items that include a directional letter, arrow, and degree heading (e.g., S ↑ 180). They must decide the direction implied by these indicators (or on one part of the test they must decide whether two of three indicators agree) or indicate that conflicting information is presented in the item. We used these two tests to identify an ability that we initially called Perceptual Speed (Complex), which we distinguished
from an ability that is identified with more traditional, simple measures of perceptual speed.

In previous predictive validity studies, both the Dial Reading Test and the Directional Headings Test have shown high validity for predicting the success of Federal Aviation Administration (FAA) air traffic controllers and for performance on our TRACON simulation task. These two tests also show a high intercorrelation ($r = .74$, as reported in Ackerman & Kanfer, 1993b). In fact, similar to the findings of the Army–Air Force with the Dial and Table Reading Test ( Guilford & Lacey, 1947), our version of the Dial Reading Test and the FAA's Directional Headings Test had mathematically the highest validity for TRACON performance in a large battery of cognitive and perceptual speed ability tests. A unit-weighted z-score composite was computed from these two measures and is referred to as Perceptual-Speed-Complex (PS-Complex).

**Personality Measures**

In order to investigate the relations between criterion task performance and the broad personality domain, the NEO Personality Inventory Revised (NEO-PIR; Costa & McCrae, 1992) was administered. The NEO-PIR is a 240-item inventory designed to assess five broad factors of personality. The five broad personality factors assessed by the NEO-PIR are Extroversion, Neuroticism, Openness, Agreeableness, and Conscientiousness. The 20-item Spielberger (1983) State-Trait Anxiety measure was also administered under trait instructions.

In addition to the five broad personality factor scales, a measure was included that attempts to bridge the personality and cognitive ability domain, namely the 59-item Typical Intellectual Engagement Scale. Typical intellectual engagement (TIE) is a personality construct that represents an individual's aversion or attraction to tasks that are intellectually taxing and is thus related to acculturative and purposeful development and expression of certain intellectual abilities (e.g., crystallized intelligence). Scores on this measure have different correlations with the broad ability classes of fluid and crystallized intelligence (with higher correlations between TIE and crystallized intelligence (Ackerman & Goff, 1994; Goff & Ackerman, 1992).

**Vocational Interest Measure**

Vocational interests were assessed with American College Testing's 90-item Unisex Edition of the ACT Interest Inventory (UNIACt; Lamb & Prediger, 1981). The UNIACt inventory is designed to assess six different factors of vocational interests (Holland, 1959, 1985) by asking respondents to indicate their interest or liking for doing particular job tasks, irrespective of their competency to do them. The six scales in the battery are Realistic, Investigative, Artistic, Social, Enterprise, and Conventional.

**Self-Ratings of Ability and Self-Concept**

For self-ratings of ability, a 31-item questionnaire was used. The items represent three broad factors of abilities: spatial-math, verbal, and self-regulation. Trainees were instructed to respond with a self-evaluation relative to other persons their age.

Self-concepts for competencies and knowledge in several specific areas were assessed with a 32-item questionnaire developed specifically for TRACON-relevant abilities and skills. The areas assessed were self-concepts of mechanical, verbal, math, spatial, clerical, and science competencies, as well as self-concepts of self-management ability and stress resistance.

**Motivational Skills**

The Motivational Skills measure was an 18-item questionnaire assessing aspects of self-confidence for learning, studying, and performing in test situations (see Ackerman & Kanfer, 1993b). The Motivational Skills questionnaire has been shown to provide incremental prediction of TRACON performance after abilities are partialled out.

**Self-Efficacy Measures**

A series of self-report questionnaires was administered before and after the TRACON simulation trials (which occurred on Sessions 3–8). Trainees
were asked to indicate their self-efficacy for components of TRACON task performance (handling arrivals and overflights) and for overall performance (aggregate planes handled and performance relative to the population of college students). Questions were administered in an ordered fashion. In the Interim Questionnaire, questions referred to performance on the current session. For example, “I can land $xx$ percentage of the planes in the upcoming session” (where $xx$ can be replaced with 20, 40, 60, 80, and 100; or “Overall, I can perform TRACON at least as well as $yyy$ other students” (where $yyy$ can be replaced with a few, some, many, most, or almost all). For the Task Perceptions Questionnaire, questions pertained only to planes handled and overall performance. These questions referred to expected performance on the next session. Consistent with previous methods used to assess self-efficacy, trainees were instructed to indicate their confidence in performing at each target level with a 0–10 scale, with 0 indicating no confidence and 10 indicating certain (“that I can do it”). Self-efficacy scores were obtained by summing confidence scores across each set of items.

**Proximal-Concomitant Measures**

In addition to self-efficacy, several other scales were administered either just before each TRACON session or just after. Prior to task performance, trainees were asked about negative motivational thoughts (e.g., “Between the last session and today, I wished that I was done with this project”) and positive motivation thoughts (e.g., “Between the last session and today, I imagined myself making no errors on the task”). Subsequent to each session, trainees were asked questions about frequency of thoughts related to planning, off-task, positive cognitions, negative affect, and motivation during the current TRACON session.

**Procedure**

The experiment was completed in eight sessions (see Figure 1) over a 2-week period. The first session was devoted to ability testing. The second session was devoted to personality and other self-report testing and to viewing an instructional videotape specifically designed for the TRACON task. The 60-min videotape featured the major task components, the rules regarding operation of the computer interface (keyboard and trackball), display characteristics, and procedures for accomplishing the controller task. Two quizzes were administered, one about halfway through the videotape and the other at the end (in order to maintain motivated attention to the task). Actual TRACON task practice started at Session 3 and proceeded through Session 8. Each of these practice sessions included five simulation scenarios. Each scenario (trial) lasted 30 min. Order of ability testing was constant across trainees, but order of simulations was counterbalanced in a Latin-square design. Breaks were given after each simulation–task trial, and self-report questionnaires were administered at the beginning and end of each TRACON practice session. Trainees were then debriefed and paid. The entire experiment was completed in 24 hr and included 6 hr of testing, 1 hr of task-based video instruction, 2 hr of accumulated breaks, and 15 hr of total time on task in TRACON.

**Results**

The presentation of the results from this study has been divided into five sections, as follows: (a) descriptive statistics of distal measures and performance measures (both overall TRACON performance and performance components); (b) distal measure (ability and nonability)—performance associations (and a description of gender differences in the context of task requirements); (c) concomitant measures, including descriptive statistics and performance associations; (d) an integrated prediction model (for performance and concomitant measures), which uses both multiple regression and path-analysis perspectives. In the last section, (e) we present a detailed sketch of the broader predictor space by means of multidimensional scaling and factor analytic techniques.

**Descriptive Statistics**

Means, standard deviations, squared multiple correlations, and intercorrelations are provided for 30 composite distal measures in Table 1.
Composites were formed by unit weighted z scores of the tests in each content set (i.e., Spatial Ability, Verbal, PS-Complex, PS-Simple, Math). For summary purposes (see below), an overall general ability estimate was also created, with a unit-weighted composite of all 14 ability test measures. The other 16 distal measures (personality, interests, self-concept, and self-ratings of ability) were summed composites of raw items. These composite measures revealed means and variances well...
within previous normative results. The self-report measures generally had quite high internal-consistency reliabilities. In fact, only two self-report measures had reliabilities below $\alpha = .80$; namely Self-Management self-concept ($\alpha = .68$) and Self-Regulation self-estimate of ability ($\alpha = .77$).

**Skill Acquisition: TRACON Data**

This section describes the effects of practice on TRACON performance. Below, the results are presented for overall performance and are followed by the results for two important components.
of overall performance, namely the handling of arrivals and overflights.\(^2\)

**Overall Performance**

Previous studies with this version of TRACON clearly demonstrated that the task is complex and, within the parameters of our simulations, quite difficult early in practice (Ackerman, 1992; Ackerman & Kanfer, 1993b). On Trial 1, out of 28 possible planes, the number of actual successful plane handles was \(M = 5.89, SD = 4.67\). Although we consider this to be good initial performance, we should note that 12% of the trainees had no successful plane handles on Trial 1, even though all trainees had already received an explicit 1-hr video instruction on the task. By the end of 30 trials of practice, overall performance was much improved \((M = 19.49, SD = 6.57)\). At the end of training, 2% of the trainees were able to handle all 28 of the planes. Mean performance levels across practice are shown in Figure 2. A repeated measures analysis of variance (ANOVA) on overall performance confirmed the significant effect of practice shown in the figure, \(F(29, 2668) = 145.03, p < .01\).

**Component Performance**

Although the simulations had different numbers of arrivals (12 per trial), departures (approximately eight per trial), and overflights (approximately eight per trial), examination of performance on each type of flight helps illustrate the difference in difficulty levels of the respective types of flights. To save space and because departures represent moderate difficulty (i.e., more difficult than overflights but less difficult than arrivals), these component measures have not been included. Mean performance levels for arrivals and overflights are shown in lower portion of Figure 2. For Trial 1, \(M_{\text{Arrivals}} = 1.69\) and \(M_{\text{Overflights}} = 2.51\). At the end of practice, substantial improvement on both types of flights was indicated: Trial 30, \(M_{\text{Arrivals}} = 7.99\) and \(M_{\text{Overflights}} = 6.51\). Although trainees were handling a greater number of arrivals than overflights at the end of practice, the different number of flight types available in each trial complicates making direct comparisons. That is, at the end of practice, trainees were handling 66% of the arrivals but 81% of available overflights.

**Distal Measure–Performance Associations**

**Ability–Performance Relations**

Although the purpose of the two quizzes administered during the videotape TRACON instructions was to maintain trainee motivation for attending to the instructional material, data from the quizzes did provide some insight into the nature of predictor-criteria relations. Ability measures well-predicted a summed measure across the two quizzes—all ability composites showed significant correlations and a multiple correlation indicated substantial ability overlap with the quizzes \((R = .59, p < .01)\). In turn, quiz scores significantly correlated with the proximal measures of negative motivational thoughts \((r = -.44, p < .01)\) and self-efficacy \((r = .33, p < .01)\). Moreover, quiz scores were highly correlated with later TRACON performance across task practice \((e.g., \tau_{\text{Session} 1} = .62; \tau_{\text{Session} 6} = .66)\).
man, 1992; Kanfer & Ackerman, 1993b), for a discussion of interrelations among performance predictors it is necessary to document fully the substantial cognitive demands of the task, when trainees (a) first confront the task and (b) have obtained substantial task practice. For these and subsequent analyses, we concentrated on macro changes in performance by averaging scores on TRACON by session of practice (five simulation trials were administered on each of the six practice sessions). In a task that is much more quickly learned, such large time slices obscure the learning process. However, for a complex task such as TRACON, this level of analysis provides additional stability for the performance scores and more manageable data for structural analysis, without sacrificing validity or obscuring the patterns of learning and skill acquisition. Although session-to-session correlations were quite high ($r_s = .85, .92, .96, .96, \text{and} .95$ for Session 1 with Session 2, Session 2 with Session 3, etc.), Session 1 correlated with Session 6 performance ($r = .77$), consistent with the general finding of simplex-like patterning of performance data over multiple occasions (e.g., see Ackerman, 1987; Humphreys, 1960). Although the latter result indicates that 60% of the individual differences variance in Session 6 performance can be explained by Session 1 performance, such a relationship also indicates that even relatively modest differences found between predictor variable correlates with Session 1 and with Session 6 measures may be statistically significant.

The analyses below focus on overall TRACON performance, as well as differences between performance components of handling arrivals and overflights. First, we describe the correlations between the ability composites and TRACON performance over practice.

**Overall performance.** Overall performance is well predicted across all sessions of practice by the general ability composite ($r_s = .49, .64, .57, .56, .58, \text{and} .56$ for Sessions 1–6, respectively). Examination of constituent content ability measures, as shown in Figure 3, indicates that only the Math and Spatial broad content abilities are substantial (mean $r = .56$ and .53, respectively) and stable predictors of performance, the Verbal content ability shows only a modest correlation with performance (mean $r = .24$). The two PS abilities show substantial validity in predicting overall TRACON performance with the PS-Complex (mean $r = .60$) showing higher validity than PS-Simple (mean $r = .44$). These results are consistent with previous task analyses and extant validity data in two previous studies (Ackerman, 1992; Ackerman & Kanfer, 1993b).

Secondary analyses can be predicated on general performance across all sessions of practice. Toward this end, a measure of overall planes handled was computed across all 30 trials and correlated with the general and the constituent-content and PS-ability composites. The correlations between abilities and overall performance across all sessions of practice are as follows: General, $r = .59$; Spatial, $r = .56$; Math, $r = .59$; Verbal, $r = .24$; PS-Complex, $r = .63$; and PS-

---

3 As shown in Table 1, PS-Complex correlates substantially with the Spatial ability composite ($r = .66$) and with PS-Simple ($r = .56$), suggesting that the ability represented by this composite falls somewhere in-between the two more traditional reference factors.
Simple, $r = .46$ (all significant $p < .05$). A single-step multiple regression (with all ability measures except for the aggregate general composite) yielded a multiple $R$ of .70 or an $R^2 = .49$. That is, taken together, the five ability composites accounted for 49% of the variance in overall TRACON performance across 15 hr of task practice.

**Arrivals and overflights.** Separating the performance components of arrivals and overflights illustrates the task analysis that described arrivals as more difficult and more demanding of spatial abilities than overflights, which are hypothesized to be more easily learned and thus relatively better predicted by PS abilities than broad content abilities. Figure 4 shows that initial performances on arrivals and overflights were equivalently well-predicted by the general ability composite. Over practice, validity coefficients for arrivals and overflights diverge, with general ability showing higher correlations with arrival performance and lower correlations with overflights.

Figure 5 shows arrival and overflight correlations for four ability composites (Verbal ability having been dropped as a result of demonstrated lack of substantial validity for predicting TRACON performance, especially relative to the other ability measures). Again, concordant with our expectations, Spatial and Math ability diverge in validity for predicting arrivals and overflights after initial practice (with greater validity for predicting arrivals). The two PS measures show greater similarity in predicting performance of both arrivals and overflights.

**Summary.** The results described above support three points about the relations between abilities and performance on TRACON, as follows:

1. Performance throughout practice is well-predicted by measures of cognitive and intellectual ability. The measures most highly associated with task performance were Spatial, Math, and PS-Complex abilities. Verbal ability showed only a modest validity for predicting performance on TRACON at any stage of practice.

2. Across practice, ability measures accounted for nearly 50% of the variance in TRACON performance.

3. Arrival and overflight components of performance showed substantial and equivalent ability demands early in practice, but with practice, spatial and math abilities were more predictive of arrival than overflight components. In contrast, PS abilities showed generally stable correlations with both arrivals and overflights across practice.

**Gender Differences**

A robust finding in the ability literature is group differences in performance on tasks that require complex spatial information processing, with men showing an overall performance advantage over women (e.g., Lohman, 1986). Our previous studies have demonstrated similar group differences in performance on TRACON, with male trainees showing a substantial mean performance advantage relative to female trainees. Because some of these differences may be determined by ability differences, by personality, and other nonability differences, it is necessary to review the TRACON performance by gender.

**TRACON Performance**

Similar to previous studies (e.g., Ackerman, 1992), there was both a significant main effect for trainee gender on overall performance, $F(1, 91) = 19.66, p < .01$, and an interaction between trainee gender and sessions of practice, $F(5, 455) = 5.49, p < .01$. The means for each group over practice are shown in the top panel of Figure 6. As shown, the interaction is indicated as a divergence in...
Correlations between ability composites and performance components (arrivals and overflights) on the Terminal Radar Approach Controller Task, over sessions of practice (five simulation trials per session).

Concordant with the earlier ability-performance analyses of arrivals and overflights, which showed higher correlations between arrival performance and math-spatial abilities than for overflights, the differences between gender groups showed greater advantages to men over women in handling arrivals. The lower panels of Figure 6 show these effects. A repeated measures ANOVA of arrivals showed a main effect of gender, $F(1, 91) = 32.25, p < .01$, and a significant interaction between gender and practice, $F(5, 455) = 7.74, p < .01$, with performance diverging between groups with practice. In contrast, a repeated measures ANOVA on overflights showed a significant, but smaller effect of gender, $F(1, 91) = 4.32, p < .05$, and a nonsignificant interaction between gender and practice, $F(5, 455) = 2.05, ns$. At the end of practice, men and women have equivalent performance on overflights ($M_{men} = 6.44, M_{women} = 6.21$). Again, these results support the assertion that handling arrivals requires more complex spatial...
information processing than does handling overflights.

Nonability Predictor–Performance Relations

**Personality.** Personality constructs have often been implicated in determinations of stress reactivity or stress resistance in performance contexts (e.g., see Vickers, 1991). We first correlated measures of the five major personality composites with overall performance on TRACON. None of the correlations reached traditional levels of significance: Neuroticism, \( r = -.15 \); Extroversion, \( r = -.02 \); Openness, \( r = -.01 \); Agreeableness, \( r = .02 \); Conscientiousness, \( r = -.03 \); and Anxiety, \( r = -.19 \). A multiple regression with these seven composites similarly yielded no significant result (\( R = .24 \) or \( R^2 = .06 \)). Simple correlations between personality composites and daily session measures of TRACON yielded similar results, as did separate correlations for arrivals and overflights. Only Anxiety revealed a significant correlation with performance and only for Session 1 (\( r = -.28, p < .01 \)).

**Vocational interests.** It has often been claimed in the vocational literature that emotional stress is generated when a mismatch exists between an individual's vocational interests and the task that the individual is asked to perform (see, e.g., Dawis & Lofquist, 1984). We computed correlations between the scales of the UNIACT and aggregate TRACON performance, with the following results: Realistic, \( r = .31 \); Investigative, \( r = .19 \); Artistic, \( r = -.14 \); Social, \( r = -.07 \); Enterprising, \( r = -.07 \); and Conventional \( r = .07 \). The correlations between the Realistic and Investigative scales and performance are the only ones that reached both statistical and practical significance, jointly accounting for about 10% of the variance in TRACON performance; \( R = .31, F(2, 90) = 4.90, p < .01 \).

**Self-ratings of ability and self-concept.** The three composite scales of ability self-ratings (SR) and the eight measures of academic ability self-concept were factor analyzed to yield three main composites. The first composite (SR1) was comprised of items from math, spatial, mechanical, and science self-concept and self-ratings of similar abilities. The second (SR2) was comprised of items from verbal self-concept and self-ratings of general verbal ability and more specific reading, vocabulary, and writing abilities. The third (SR3) was comprised of items from self-management, clerical, and stress-resistant self-concept, and self-ratings of self-control and coping abilities. Correlations between these three composites and TRACON performance indicated that the first composite was highly and consistently related to performance (mean \( r = .46 \)), and the remaining two composites were essentially unrelated to TRACON performance (SR2: mean \( r = -.08 \); SR3: mean \( r = .12 \)). These results indicate that selected self-ratings of ability and academic ability self-concept are effective predictors of performance in a complex task. In the aggregate, these measures accounted for 24% of the overall variance in TRACON; \( R = .49, F(3, 87) = 9.10, p < .01 \). In particular, the ratings of math, spatial, and mechanical abilities accounted for about 20% of the variance in TRACON performance across practice sessions.

**Motivational skills.** As in previous studies (Ackerman & Kanfer, 1993b), an 18-item measure of Motivational Skills was included in this study. Consistent with results from predicting success in air traffic control tasks (both in the laboratory and in the field with FAA trainees), the Motivational Skills measure showed consistent, modest, and significant correlations with performance across all sessions of TRACON performance (mean \( r = .25 \)). For aggregate performance, the correlation with Motivational Skills was \( r = .27, p < .01 \), or about 7% of the variance.

**Proximal Measures**

In addition to the sets of distal individual differences measures administered prior to trainees' exposure to TRACON, an additional set of proximal measures administered during the training sessions were included. These measures included measures of ego involvement, stress reactivity, and task performance. The measures were administered at the end of each practice session to assess the trainees' performance and to determine the effects of training on their ability to perform the task.

*Several authors (e.g., Snow, 1989) have argued for nonlinear relations between personality measures and ability measures and by implication between personality and task performance measures. We examined quadratic and cubic regressions between personality and TRACON performance measures and did not find any significant increment in associations beyond the linear regressions. However, it is possible that because our measures were geared to a normal sample, such effects may not generalize to a sample of tests that are better geared to assessment of psychopathology and/or to trainees that are more broadly sampled from the population at large.*
mal (and concomitant) measures was administered. Specifically, prior to each session of TRACON practice, the Interim Questionnaire (16 items) was completed by the trainees. Similarly, at the end of each session of TRACON practice, a Task Perceptions Questionnaire (44 items) was completed by the trainees. Of critical importance is the fact that the first Interim Questionnaire was administered prior to any direct TRACON experience but 1 or 2 days after the video instruction on TRACON was presented to the trainees. The first interim questionnaire captures the proximal anticipations of trainees for their initial confrontation with TRACON.

Interim questionnaire. Two subscales from the Interim Questionnaire were of critical importance for the current investigation: negative motivation and positive motivation thoughts. The Negative scale is composed of six items such as “I wished that I was done with this project” and “I got discouraged when I thought about today’s session.” The Positive scale is composed of seven items such as “I imagined myself doing extremely well on today’s session” and “I made a specific plan for how to perform some part of the task.” Both measures significantly predicted initial performance on TRACON Session 1, negative motivation \( r = -0.29 \) and \( -0.38 \) and positive motivation \( r = 0.26 \) and \( 0.22 \) for arrivals and departures, respectively. In addition, the negative motivation items administered prior to task engagement showed stability in predicting TRACON performance over the subsequent days of task practice (see Figure 7). The positive motivation thoughts only showed stable correlations with the overflight component of TRACON.

Self-efficacy prior to task engagement. At the end of the Interim Questionnaire, trainees were asked to indicate their self-efficacy for components of TRACON task performance (handling arrivals and overflights) and for overall performance (aggregate planes handled relative to the population of college students). As with the negative motivation and positive motivation items, these measures were administered 1 to 2 days after trainees had viewed the 1-hr instructional video on TRACON, but just before the trainees had any direct experience in performing TRACON.

All four of the self-efficacy measures correlated significantly with initial TRACON performance (\( \text{SE}_{\text{Arrivals}} r = 0.23; \) \( \text{SE}_{\text{Overflights}} r = 0.32; \) \( \text{SE}_{\text{Planes handled}} r = 0.26; \) and \( \text{SE}_{\text{Performance}} r = 0.38 \)). These measures were also highly intercorrelated, so that an aggregate measure of self-efficacy correlated \( r = 0.32 \) with initial TRACON performance. These correlations increased after the first session and were stable throughout the remaining sessions. For example, the composite self-efficacy measure correlated \( r = 0.40, 0.38, 0.39, 0.42, \) and \( 0.43 \) with performance on the respective five remaining sessions of TRACON performance. The self-efficacy measure that showed the highest individual validity was the performance measure, which correlated \( r = 0.48 \) with overall performance on TRACON, accounting for 23% of the variance in performance across practice sessions.

In summary, measures of proximal thoughts and self-efficacy prior to engagement in TRACON performance showed validity for predicting individual differences in task performance. Specifically, greater incidence of negative motivation thoughts was significantly and substantially associated with lower TRACON performance throughout practice, and to a somewhat smaller degree, greater incidence of positive motivation thoughts was significantly associated with positive TRACON performance throughout practice. Also,
higher levels of self-efficacy (especially as perceived in comparison with other college students) were associated with TRACON performance.

**Concomitant Variables and Consequences of Practice: Measures Taken During Skill Acquisition**

**Negative Motivation and Positive Motivation**

Repeated measures ANOVAs on the negative motivation and positive motivation scales revealed significant increases for both negative motivation thoughts, $F(5, 455) = 4.02$, and positive motivation thoughts, $F(5, 455) = 9.92$, over sessions of practice (both $p < .01$). However, the negative motivation scale showed a main effect for gender, $F(1, 91) = 3.95$, $p < .01$—with women reporting higher frequency of negative motivation thoughts. In addition, a significant interaction of Gender x Session, $F(5, 455) = 3.08$, $p < .01$, was obtained. This reflected a convergence of male and female reporting of negative motivation thoughts over practice. No gender main or interaction effects were found for positive motivation.

Incidence of negative motivation or positive motivation thoughts, subsequent to the initial TRACON practice session, showed diminished validity for predicting TRACON performance. In the aggregate, these results suggest that trainees' initial meta-motivational strategies for task learning have important and persistent influences on task performance.

**Task Perceptions Questionnaires**

In contrast to the Interim Questionnaires (which were administered prior to each TRACON session), Task Perceptions Questionnaires were administered immediately following the last simulation for each of the six TRACON sessions. These measures concerned the frequency of thoughts that trainees had during the TRACON session. Measures of planning (e.g., “I thought about how many more planes I could handle”), off-task (e.g., “I daydreamed while doing the task”), positive cognitions (e.g., “I thought about ways to excel on this task”), negative affect (e.g., “I got upset when I made a mistake”), and motivation (“I gave the task my maximum effort”) were included. Repeated measures ANOVAs were performed for each of these measures with gender as a between-subject variable and practice as a within-subject variable. These ANOVAs are presented in Table 2. Results indicate several significant differences between gender groups, over practice, and interactions between two variables.

For frequency of planning thoughts, there was no main effect of gender, but significant increases in planning over the course of practice and a significant interaction of Gender x Practice. Examination of the means showed that the interaction was the result of women reporting less planning than men early in task practice but more planning later in task practice.

For off-task thoughts and negative affect thoughts, only a main effect of practice was significant. The means showed that although Session 1 off-task and negative-affect thought frequency was high (at the point where performance was lowest), the frequency dropped substantially on Session 2. Off-task thoughts showed a slow increase through later sessions of practice, whereas negative affect thoughts remained stable after Session 2. These results are consistent with earlier studies (e.g., Kanfer & Ackerman, 1989), that indicate off-task thoughts are often indicative of problems in negative affective responses to stressful task demands.

### Table 2

**Repeated Measures Analysis of Variance for Task Perceptions Measures Over TRACON Practice Sessions**

<table>
<thead>
<tr>
<th>Task perceptions measure</th>
<th>Planning</th>
<th>Off-task</th>
<th>Positive cognitions</th>
<th>Negative affect</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$df$</td>
<td>$MSE$</td>
<td>$F$</td>
<td>$MSE$</td>
<td>$F$</td>
</tr>
<tr>
<td>Gender</td>
<td>1, 90</td>
<td>669.46</td>
<td>0.20</td>
<td>379.06</td>
<td>0.43</td>
</tr>
<tr>
<td>Sessions of practice</td>
<td>5, 450</td>
<td>43.52</td>
<td>14.12**</td>
<td>44.94</td>
<td>13.92**</td>
</tr>
<tr>
<td>Gender x Sessions of practice</td>
<td>5, 455</td>
<td>43.52</td>
<td>3.14**</td>
<td>44.94</td>
<td>1.25</td>
</tr>
</tbody>
</table>

Note. TRACON = Terminal Radar Approach Controller.

**$p < .01$.**


that is, emotion control problems (see Kanfer & Ackerman, 1995).

**Interim Self-Efficacy Measures**

Four measures of self-efficacy for TRACON performance were administered prior to each TRACON session. These pertained to self-efficacy for (a) percentage of planes handled, (b) performance (relative to other students), (c) overflights, and (d) arrivals. Mean self-efficacy scores for each of the four measures over Practice × Gender are presented in Figure 8. Repeated measures ANOVAs for each are shown in Table 3. For each of the four self-efficacy measures, there were substantial and significant main effects of gender on scores (with women having lower mean self-efficacy for all aspects of the TRACON task), and sessions of practice (with self-efficacy dropping from Session 1 to Session 2 and then rising after Session 2). Significant Gender × Session of Practice interactions were found for SE_handled and SE_arrivals, with a divergence of scores for men and women and no significant interactions for SE_performance or SE_overflights.

It is interesting that even after five sessions of TRACON practice, the final SE_performance and SE_arrivals ratings by women did not exceed those of men at Session 1, even though mean overall performance on TRACON showed a substantial advantage by women at Session 6 vs. men at Session 1.

Because self-efficacy has been discussed as a potentially causal construct for predicting task performance and repeated measures of self-efficacy were obtained during the course of task practice, this study provides a means for evaluating the changes in predictive validity for self-efficacy as trainees gained experience and skill over the course of the six TRACON sessions. Figure 9 represents several correlational perspectives toward evaluating the predictive validity of self-efficacy across practice sessions for each of the four presession measures of self-efficacy (i.e., SE_handled, SE_performance, SE_arrivals, and SE_overflights). The first curve (Day 1) shows the correlations between self-efficacy, as measured prior to any direct experience with performance on TRACON and performance on TRACON across the six practice sessions. Overall TRACON performance is the criterion variable for SE_handled and SE_performance; for SE_arrivals, the criterion is arrivals handled; and for SE_overflights, the criterion is overflights handled). As shown in

![Figure 8. Mean self-efficacy (SE) expectation composite scores by gender over TRACON Terminal Radar Approach Controller Task practice. Measures administered prior to each TRACON session.](image)
Table 3
Repeatead Measures Analysis of Variance for Self-Efficacy (SE) Measures Over TRACON Practice Sessions

<table>
<thead>
<tr>
<th>Variable</th>
<th>df</th>
<th>MSE</th>
<th>F</th>
<th>df</th>
<th>MSE</th>
<th>F</th>
<th>df</th>
<th>MSE</th>
<th>F</th>
<th>df</th>
<th>MSE</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>1, 91</td>
<td>214.66</td>
<td>29.14***</td>
<td>289.46</td>
<td>31.92**</td>
<td>331.36</td>
<td>15.39***</td>
<td>292.92</td>
<td>32.46**</td>
<td>1, 91</td>
<td>189.29</td>
<td>3.95*</td>
</tr>
<tr>
<td>Sessions of practice</td>
<td>5, 455</td>
<td>21.70</td>
<td>61.73***</td>
<td>23.64</td>
<td>13.92**</td>
<td>27.37</td>
<td>32.02**</td>
<td>25.91</td>
<td>69.58**</td>
<td>5, 455</td>
<td>15.92</td>
<td>4.02**</td>
</tr>
<tr>
<td>Gender × Sessions of practice</td>
<td>5, 455</td>
<td>21.70</td>
<td>3.23***</td>
<td>23.64</td>
<td>0.70</td>
<td>27.37</td>
<td>1.14</td>
<td>25.91</td>
<td>7.06**</td>
<td>5, 455</td>
<td>15.92</td>
<td>3.08**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task perceptions questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>1, 91</td>
<td>317.59</td>
<td>33.13***</td>
<td>258.28</td>
<td>30.29**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sessions of practice</td>
<td>5, 455</td>
<td>18.48</td>
<td>26.08***</td>
<td>21.26</td>
<td>99.33**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender × Sessions of Practice</td>
<td>5, 455</td>
<td>18.48</td>
<td>0.74</td>
<td>21.26</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. TRACON = Terminal Radar Approach Controller; SE = self-efficacy.
*p < .05. **p < .01.

Figure 9. Correlations between four self-efficacy (SE) measures and performance on the Terminal Radar Approach Controller Task (TRACON). Curves represent (a) correlations between pretask (Day 1) self-efficacy and TRACON performance over practice sessions; (b) synchronous (same session) correlations between self-efficacy and TRACON performance; (c) lagged correlations (self-efficacy with previous session TRACON performance); and (d) partial correlations (synchronous correlations with previous session TRACON performance partialed out).
Figure 9, the validity coefficients for self-efficacy (Day 1) and subsequent performance were generally stable across TRACON practice, though smallest for the SE_{Arrivals} and largest for SE_{Performance} measures.

The second curve (synchronous) shows the correlations between SE measures administered just prior to each day's session and performance on TRACON for that session. These correlations increase substantially as practice on TRACON proceeds and is most pronounced for SE_{Arrivals} and least pronounced for SE_{Performance}. The third curve (lagged) shows the correlations between each session's SE measure (except for the measure administered in Session 1) and the previous session TRACON performance. For all of the SE measures, these correlations are either larger than the synchronous correlations or are equivalent to the synchronous correlations, indicating that any session's SE measures may have more in common, or at least as much in common, with the previous day's performance level (a retrospective view) than with the upcoming performance (a prospective view). Indeed, if the previous session's performance is partialed out of the synchronous correlations (the fourth curve: partial), it appears clear that prior performance accounts for nearly all of the shared variance between SE estimates and TRACON performance. The only general exception to this finding was for Session 3, where significant positive partial correlations are found for SE_{Handled} and marginally significant (p = .09) correlations were found for SE_{Performance} and (p = .06) for SE_{Overflights}. That is, self-efficacy, as measured prior to Session 3, was found to account for a significant though small, mean (r = .22, r^2 = .05) amount of Session 3 TRACON performance variance, not explained by prior task performance. However, at the end of TRACON practice, SE measures showed no significant correlations with task performance, once prior session performance variance was partialed out. Although it is not possible to specifically determine the test–retest reliabilities for the self-efficacy measures given the changing nature of performance and self-efficacy during training, it is important to note that a correction for attenuation of the raw correlations for self-efficacy measures and TRACON performance would lead to slightly larger partial correlations (e.g., see Stouffer, 1936, regarding the interpretation of partial correlations based on measures with imperfect reliability).

**Task Perceptions Self-Efficacy Measures**

Two measures of self-efficacy for TRACON performance were administered immediately after each TRACON session. These pertained to self-efficacy for (a) percentage of planes handled and (b) performance (relative to other students) in a format that was identical to the Interim Questionnaire measures. Mean self-efficacy scores for each of the two measures over practice by gender are presented in Figure 10. Repeated measures ANOVAs for each are shown in the top part of Table 3. These measures show similar patterns to those from the Interim Questionnaire, with the exception of the absence of a drop in scores from Session 1 to Session 2. The obvious reason for this difference is that the Session 1 self-efficacy measures were administered after the trainees had their initial experience with TRACON and the first Interim Questionnaire was administered prior to actual performance experience. Significant main effects for gender and for sessions of practice were found for both measures, but no significant interactions were found for either measure. Such results indicate general improvement in self-efficacy with practice and persistently higher self-efficacy scores for men.

**Integrated Prediction Model for Performance and Concomitant Variables: Multiple Regression and Correlation Results**

**Performance**

As in previous studies with TRACON (e.g., Ackerman, 1992; Ackerman & Kanfer, 1993b), raw correlations between various predictors and task performance only give partial information, in that
Common variance among predictor variables serves to limit the overall predictive validity of an integrated battery of tests. Multiple correlation and regression procedures allow for the determination of the incremental validity of each set of predictor variables in a general equation for predicting task performance. The general logic of the analyses in this section was first to enter variables that are assumed to be causal antecedents. That is, ability measures were entered into the prediction equation and then followed by the other distal variables (self-concept and vocational interests) and proximal variables (motivational thoughts and self-efficacy). Finally, gender was added to the equation to examine whether gender differences accounted for variance in performance after all the predictor measures had been taken into account. However, only measures obtained prior to task performance were included in the prediction equation. We dropped the personality measures from these computations because they did not provide significant raw correlations with TRACON performance.

Table 4
Hierarchical Multiple Regression and Correlation Results for Performance

<table>
<thead>
<tr>
<th>Step</th>
<th>Increment in $R^2$</th>
<th>$F$ to add</th>
<th>$df$</th>
<th>Increment in $R^2$</th>
<th>$F$ to add</th>
<th>$df$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prediction of overall TRACON performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.403</td>
<td>.020</td>
<td>.90</td>
<td>3, 82</td>
<td>.458</td>
<td>.018</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.410</td>
<td>.007</td>
<td>.47</td>
<td>2, 80</td>
<td>.466</td>
<td>.008</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>.457</td>
<td>.046</td>
<td>3.31*</td>
<td>2, 78</td>
<td>.540</td>
<td>.074</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>.457</td>
<td>.000</td>
<td>.01</td>
<td>1, 77</td>
<td>.547</td>
<td>.007</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.467</td>
<td>.010</td>
<td>1.46</td>
<td>1, 76</td>
<td>.568</td>
<td>.021</td>
</tr>
<tr>
<td>Total</td>
<td>.467</td>
<td>NA</td>
<td>4.75**</td>
<td>14, 76</td>
<td>.568</td>
<td>NA</td>
</tr>
<tr>
<td>Prediction of overflight component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ability</td>
<td>.331</td>
<td>.321</td>
<td>8.41**</td>
<td>5, 85</td>
<td>.226</td>
<td>.226</td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.340</td>
<td>.009</td>
<td>.38</td>
<td>3, 82</td>
<td>.236</td>
<td>.010</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.347</td>
<td>.007</td>
<td>.42</td>
<td>2, 80</td>
<td>.240</td>
<td>.004</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>.390</td>
<td>.043</td>
<td>2.75</td>
<td>2, 78</td>
<td>.328</td>
<td>.088</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>.392</td>
<td>.002</td>
<td>.39</td>
<td>1, 77</td>
<td>.328</td>
<td>.000</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.392</td>
<td>.000</td>
<td>.00</td>
<td>1, 76</td>
<td>.328</td>
<td>.000</td>
</tr>
<tr>
<td>Total</td>
<td>.392</td>
<td>NA</td>
<td>3.51**</td>
<td>14, 76</td>
<td>.328</td>
<td>NA</td>
</tr>
<tr>
<td>Prediction of arrival component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ability</td>
<td>.317</td>
<td>.317</td>
<td>7.89**</td>
<td>5, 85</td>
<td>.447</td>
<td>.447</td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.351</td>
<td>.034</td>
<td>1.44</td>
<td>3, 82</td>
<td>.464</td>
<td>.017</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.378</td>
<td>.027</td>
<td>1.71</td>
<td>2, 80</td>
<td>.475</td>
<td>.011</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>.417</td>
<td>.039</td>
<td>2.61</td>
<td>2, 78</td>
<td>.505</td>
<td>.030</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>.420</td>
<td>.003</td>
<td>.51</td>
<td>1, 77</td>
<td>.515</td>
<td>.010</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.457</td>
<td>.037</td>
<td>5.13*</td>
<td>1, 76</td>
<td>.565</td>
<td>.050</td>
</tr>
<tr>
<td>Total</td>
<td>.457</td>
<td>NA</td>
<td>4.57**</td>
<td>14, 76</td>
<td>.565</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note. TRACON = Terminal Radar Approach Controller Task; NA = not applicable.
*p < .05. **p < .01.
late in practice, and an increase in the amount of variance predicted for the arrival component over practice (from 45.7% to 56.5%). Cognitive and PS abilities, admittedly the first variables entered into the equation, also accounted for the largest amount of individual differences variance in overall TRACON task performance, as well as respective overflight and arrival components of the TRACON task across extensive practice. Of the remaining predictor measures, only proximal negative and positive motivational thoughts consistently provided incremental validity, which was more pronounced in the overflight component than the arrival task component. Finally, gender accounted for a significant incremental amount of variance in the arrival component of task performance but accounted for virtually no incremental variance in the overflight component.

Concomitant measures. Considerable speculation appears in the literature about the determinants of individual differences in task self-efficacy prior to task engagement and subsequent to task practice (e.g., see Bandura, 1986; Feltz, 1982; Gist & Mitchell, 1992; Mitchell et al., 1994). We computed hierarchical regressions and correlations for initial and final self-efficacy with the same order of variable entry that was used for predicting performance. We have presented here only the results for the self-efficacy measure that pertained to overall performance, although similar results were obtained with the other three self-efficacy measures. As shown in Table 5, distal measures accounted for 31.3% of the variance in initial self-efficacy (i.e., prior to task engagement) with ability (22.8%) and self-concept (8.3%) each accounting for significant amounts. The proximal measures of

Table 5

Hierarchical Multiple Regression and Correlation Results for Proximal-Concomitant Measures

<table>
<thead>
<tr>
<th>Step</th>
<th>Session 1</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Session 6</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R²</td>
<td>Increment in R²</td>
<td>F to add</td>
<td>df</td>
<td></td>
<td>R²</td>
<td>Increment in R²</td>
<td>F to add</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prediction of self-efficacy (performance)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ability</td>
<td>.228</td>
<td>.228</td>
<td>.053**</td>
<td>5.85</td>
<td>5</td>
<td>.212</td>
<td>.212</td>
<td>.465**</td>
<td>5</td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.311</td>
<td>.083</td>
<td>3.28*</td>
<td>3.82</td>
<td>3</td>
<td>.297</td>
<td>.085</td>
<td>3.31**</td>
<td>3</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.313</td>
<td>.002</td>
<td>.13</td>
<td>2.80</td>
<td>2</td>
<td>.343</td>
<td>.046</td>
<td>2.80</td>
<td>2</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>.472</td>
<td>.159</td>
<td>11.68**</td>
<td>2.78</td>
<td>2</td>
<td>.375</td>
<td>.032</td>
<td>2.02</td>
<td>2</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.77</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.515</td>
<td>.043</td>
<td>6.90*</td>
<td>1.77</td>
<td>1</td>
<td>.500</td>
<td>.026</td>
<td>3.99*</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>.515</td>
<td>NA</td>
<td>6.29</td>
<td>13.77</td>
<td>1</td>
<td>.500</td>
<td>NA</td>
<td>2.94**</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Prediction of negative motivational thoughts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ability</td>
<td>.190</td>
<td>.190</td>
<td>4.00**</td>
<td>5.85</td>
<td>5</td>
<td>.132</td>
<td>.132</td>
<td>2.58*</td>
<td>5</td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.296</td>
<td>.106</td>
<td>4.12**</td>
<td>3.82</td>
<td>3</td>
<td>.272</td>
<td>.140</td>
<td>5.20**</td>
<td>3</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.347</td>
<td>.051</td>
<td>3.11*</td>
<td>2.80</td>
<td>2</td>
<td>.283</td>
<td>.011</td>
<td>0.58</td>
<td>2</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.77</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.77</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.360</td>
<td>.013</td>
<td>1.57</td>
<td>1.79</td>
<td>1</td>
<td>.351</td>
<td>.015</td>
<td>1.76</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>.360</td>
<td>NA</td>
<td>4.04**</td>
<td>11.79</td>
<td>1</td>
<td>.351</td>
<td>NA</td>
<td>2.94**</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Prediction of positive motivational thoughts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ability</td>
<td>.131</td>
<td>.131</td>
<td>2.56*</td>
<td>5.85</td>
<td>5</td>
<td>.194</td>
<td>.194</td>
<td>4.08**</td>
<td>5</td>
</tr>
<tr>
<td>2. Self-concept</td>
<td>.279</td>
<td>.148</td>
<td>5.62**</td>
<td>3.82</td>
<td>3</td>
<td>.306</td>
<td>.112</td>
<td>4.42**</td>
<td>3</td>
</tr>
<tr>
<td>3. Interests</td>
<td>.300</td>
<td>.021</td>
<td>1.19</td>
<td>2.80</td>
<td>2</td>
<td>.311</td>
<td>.005</td>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>4. Motivational thoughts</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.77</td>
</tr>
<tr>
<td>5. Self-efficacy</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.77</td>
</tr>
<tr>
<td>6. Gender</td>
<td>.306</td>
<td>.006</td>
<td>.69</td>
<td>1.79</td>
<td>1</td>
<td>.525</td>
<td>.027</td>
<td>4.38*</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>.306</td>
<td>NA</td>
<td>3.17**</td>
<td>11.79</td>
<td>1</td>
<td>.525</td>
<td>NA</td>
<td>6.01**</td>
<td>14</td>
</tr>
</tbody>
</table>

Note. NA = not applicable.
*p < .05. **p < .01.
motivational thoughts also accounted for a significant and substantial amount of variance (15.9%). Initially, it may appear that the covariance between the motivational thought measures and self-efficacy was the result of their being measured on the same questionnaire (self-efficacy was at the end). However, the fact that mitigates against this explanation is that the motivational thought questions were retrospective (i.e., thoughts you have had over the past 1-to-2 days), and the self-efficacy measure was prospective (i.e., your confidence in performing during the upcoming session). Gender accounted for an additional 4% of variance in task self-efficacy even after the inclusion of the distal and proximal measures.

Self-efficacy prior to the last session of practice (Session 6) was also well predicted by distal measures and proximal measures that were administered prior to task engagement, even though self-efficacy for Session 1 and Session 6 were not extremely highly correlated ($r = .52, p < .01$). In fact, ability, self-concept, and interests accounted for similar amounts of both Session 1 and Session 6 self-efficacy variance. However, initial motivational thoughts added little to the prediction of Session 6 self-efficacy, whereas Session 1 self-efficacy added a significant incremental prediction (10% variance accounted for) beyond the distal measures first included in the prediction equation. Such results are consistent with an interpretation of self-efficacy. That is, there are traitlike properties not univocally influenced by short-term experiences, and, for at least some trainees, self-efficacy may have been resilient even in the face of failure. Finally, gender accounted for an additional 3% of variance in self-task self-efficacy even after the inclusion of the distal and proximal measures.

Hierarchical regression analyses were applied to Session 1 and Session 6 motivational thought variables (also shown in Table 5). These analyses show that—similar to the prediction of self-efficacy—distal measures (i.e., ability and self-concept) account for about 30% of the variance in motivational thought frequency at Session 1 as well as at Session 6. However, at Session 6, a significant amount of variance (18%) in positive motivational thoughts was accounted for by Session 1 positive motivational thoughts. Gender provided only incremental prediction validity for positive motivational thoughts and then only at Session 6 (3%).

Path Analyses

Use of hierarchical multiple regression, as described above, provides a partitioning of variance accounted for in the dependent variable, which is partly determined by the order in which variables are entered into the equation. Even though we argue that entry of distal variables first and proximal measures afterward is the only ordering that makes logical sense, alternate perspectives might also be fruitful, for example, where interests are entered prior to self-concept. One way to estimate the relative contribution of the distal and proximal measures to predicting performance is to produce by means of LISREL or other modeling techniques a path model that simultaneously estimates all path coefficients. We conducted a series of tests that attempted to provide a model for predicting performance, which was based on the variables described in the regression section above. In performing these tests, we found that it was useful to create aggregate measures for some variables (e.g., ability and self concept). Also, several variables were dropped from the model (e.g., interests) because of findings of nonsignificant paths. The models we derived are shown in Figure 11.

![Path Analyses](image-url)
models showed reasonable fit statistics according to standard methods for assessing adequacy of models. The model for Session 1 TRACON performance, shown in the upper panel of Figure 11, yielded \( \chi^2(7) = 8.5, p = .29 \), General Fit Index (GFI) = .97, Normed Fit Index (NFI) = .95, and Non-Normed Fit Index (NNFI) = .98. An identical model when applied to Session 6 TRACON performance (lower panel of Figure 11) yielded \( \chi^2(7) = 12.32, p = .09 \), GFI = .96, NFI = .93, and NNFI = .93. (See Bentler & Bonett, 1980, for discussion of these fit statistics and their interpretation.)

Figure 11 shows the coefficients for all significant paths in the model. Of specific interest, we found similar path coefficients for ability (g) to TRACON performance (.48 for Session 1 and .52 for Session 6). In addition, a significant path was found for ability to self-concept. From self-concept, significant paths were found to both negative and positive motivational thoughts, as well as a direct path to task self-efficacy, suggesting that self-concept serves as a mediator for the association between ability and self-efficacy because there is no significant direct path from ability to self-efficacy. In addition, a significant path was found from negative motivational thoughts to TRACON performance—which is consistent with and generally supportive of the role that deficits in emotion control have on task performance (e.g., Kanfer & Ackerman, 1990, 1995). When viewed in the context of the hierarchical regressions reported earlier, these results suggest that individual differences in ability account for the major share of predicted individual differences in TRACON performance across practice, but that self-report measures, especially motivational thoughts as mediated by self-concept, also account for a nontrivial amount of variance in both initial and final TRACON performance. Self-efficacy, as assessed prior to task engagement, provides no significant contribution to the prediction of individual differences in performance that is not in turn accounted for by other measures (e.g., self-concept and motivational thought frequency).

**Analysis of the Predictor Space**

On the basis of the wide array of distal measures we administered in this study, the discussion in the literature on the general nature of the predictor space (e.g., Kanfer, Ackerman, Murtha, & Goff, 1995) and the above findings that are indicative of communality among several families of predictor measures, we decided to focus more specifically on the predictor space. We performed two different types of analyses that served to illuminate communalties across the domains of ability, personality, vocational interests, self-concept, and self-estimates of ability. The first analysis was based on a radex-type approach (for examples, see Ackerman, 1988; Marshalek, Lohman, & Snow, 1983) and used a multidimensional scaling (MDS) technique. The second was a more traditional factor analysis approach. Each is discussed in turn below.

**MDS**

Although personality variables did not show individual or incremental validity for predicting TRACON performance, there is a broader interest in evaluating the communality of personality and other distal constructs, such as ability and vocational interests. Thus we included personality with the other distal measures for this analysis. In keeping with the radex approach, we had correlations serve as proxies for similarity estimates. Intercorrelations among the following variables were computed: personality (7 measures), vocational interests (6 measures), self-concept (8 measures), self-ratings of ability (3 measures), ability (5 composite measures), and Motivational Skills (1 measure) for a total of 30 variables. (For correlations among these variables, see Table 1.) Given that some of the correlations were negative (e.g., among personality measures), a constant (1) was added to all correlations. The matrix was then subjected to KYST-3 MDS (Kruskal, Young, & Seery, 1973), a two-dimensional solution was extracted (Stress Formula 1 = .20), and, as is customary, the solution (see Figure 12) was rotated to a principal-components orientation.

There are many interesting similarities and differences in Figure 12, but we have focused on a few similarities that we found especially salient. First, the MDS solution recovered the Holland hexagon structure typically found with vocational interest measures (and is explicitly expected from the UNIACT measures). There are several close proximities among interest variables and personality measures (e.g., Social Interests and Extroversion; Artistic Interests and Openness; Conventional In-
Figure 12. KYST-3 two-dimensional solution to distal measures. Spat = spatial; Mot = motivational; Reg = regulation; PS = perceptual speed; Mech = mechanical; TIE = Typical Intellectual Engagement; SC = self-concept.

Factor Analysis

The 30-measure intercorrelation matrix was also subjected to factor analysis. Although the MDS yields meaningful distances between variables, the goal of factor analysis is to estimate a small set of hypothetical constructs that underlie the common variance among the measures. For the factor analysis, a parallel analysis criterion (Humphreys & Montanelli, 1975) was used for the determination of the number of factors to extract. The parallel analysis indicated that eight factors should be extracted. A principal axis factor analysis was used, with iterated communalities, and rotated to a varimax criterion. Oblique rotations were also performed, but they provided neither an increase in clarity nor substantial positive manifold, which thus ruled out the usefulness of a hierarchical rotation of factors. The varimax solution is presented in Table 6.

Two of the factors appear to be uniquely identified within a narrow domain. Factor 3 is identified mainly with ability (General Ability), though math self-concept also loads on the factor, and Factor 7 (Neuroticism/Anxiety) is identified with only the two respective personality measures. In contrast, Factor 1 captures a range of self-concept variables (Mechanical, Math, Spatial, and Science) and strong loadings of Realistic and Investigative vocational interests. Thus we identified Factor 1 as Science/Mechanical Self-Concept and Interests. Factor 2 captured a range of variables, including Self-Regulation self-ratings of ability and Self-Management self-concept. In addition, the personality trait of Conscientiousness, and the vocational interest traits of Enterprising and Conventional also had substantial loadings on this factor. Thus we identified Factor 2 as a Self-Management factor. Factor 4 picked up several cultural and verbal-related traits, including Verbal Ability, Verbal Self-Concept, TIE, and Openness personality traits, and Artistic interests. Thus we identified Factor 4 as “Intellectence” (after Welsh, 1991). In comparison to Factor 4, Factor 5 was contrastive, in that positive loadings were found for Verbal self-ratings of ability, Verbal Self-Concept, and Verbal Ability, but negative loadings of Conventional interests and negative loadings of Math Self-Concept. In addition, a nearly salient loading was found for the measure of Motivational Skills. By contrasting the salient loadings between Factor 4 and Factor 5, it seems that Factor 5 taps a Self-Presentation of Verbal Ability more so than actual verbal ability. The distinction between these two factors is supported by an examination of the pattern of raw correlations, as well as the MDS results reported above, where Verbal Self-Concept and Verbal self-ratings of ability were generally somewhat removed from objective measures of Verbal ability. Factor 6 is clearly an Extroversion factor, and it has substantial loadings of Social and Enterprising vocational interests as well. Factor 8 captures mostly Spatial competence, including Spatial Ability and Spatial Self-Concept, though it also has a surprising positive loading from Agreeableness.

Taken together, the MDS and factor analysis results clearly point to overlap among various trait
Table 6
Factor Loadings for Measures and Composites: Varimax Rotation

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
<th>Factor 6</th>
<th>Factor 7</th>
<th>Factor 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>-.158</td>
<td>-.216</td>
<td>-.085</td>
<td>.137</td>
<td>.058</td>
<td>-.093</td>
<td>.807</td>
<td>-.041</td>
</tr>
<tr>
<td>Extroversion</td>
<td>-.105</td>
<td>-.065</td>
<td>.048</td>
<td>.139</td>
<td>.091</td>
<td>.798</td>
<td>-.223</td>
<td>.088</td>
</tr>
<tr>
<td>Openness</td>
<td>.095</td>
<td>-.157</td>
<td>.042</td>
<td>.735</td>
<td>.150</td>
<td>.292</td>
<td>.047</td>
<td>.109</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-.042</td>
<td>-.068</td>
<td>.029</td>
<td>-.215</td>
<td>.145</td>
<td>.086</td>
<td>-.194</td>
<td>.400</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.108</td>
<td>.742</td>
<td>-.032</td>
<td>-.061</td>
<td>.004</td>
<td>-.071</td>
<td>-.212</td>
<td>.081</td>
</tr>
<tr>
<td>TIE</td>
<td>.108</td>
<td>.125</td>
<td>.210</td>
<td>.706</td>
<td>.181</td>
<td>.111</td>
<td>-.061</td>
<td>-.087</td>
</tr>
<tr>
<td>Anxiety</td>
<td>-.108</td>
<td>-.163</td>
<td>-.100</td>
<td>.015</td>
<td>.077</td>
<td>-.158</td>
<td>.877</td>
<td>-.121</td>
</tr>
<tr>
<td>UNIACT Interests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Realistic</td>
<td>.666</td>
<td>.116</td>
<td>.127</td>
<td>.188</td>
<td>-.222</td>
<td>.026</td>
<td>-.042</td>
<td>.041</td>
</tr>
<tr>
<td>Investigative</td>
<td>.639</td>
<td>-.129</td>
<td>.124</td>
<td>.386</td>
<td>-.052</td>
<td>-.132</td>
<td>-.057</td>
<td>-.093</td>
</tr>
<tr>
<td>Artistic</td>
<td>-.067</td>
<td>-.246</td>
<td>-.117</td>
<td>.702</td>
<td>.114</td>
<td>.124</td>
<td>.182</td>
<td>.255</td>
</tr>
<tr>
<td>Social</td>
<td>.036</td>
<td>-.036</td>
<td>-.062</td>
<td>.282</td>
<td>.106</td>
<td>.621</td>
<td>-.099</td>
<td>.057</td>
</tr>
<tr>
<td>Enterprising</td>
<td>-.199</td>
<td>.470</td>
<td>-.030</td>
<td>-.090</td>
<td>-.094</td>
<td>.631</td>
<td>.092</td>
<td>-.126</td>
</tr>
<tr>
<td>Conventional</td>
<td>.199</td>
<td>.576</td>
<td>.103</td>
<td>-.270</td>
<td>-.447</td>
<td>.271</td>
<td>.243</td>
<td>-.257</td>
</tr>
<tr>
<td>Self-Concept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanical</td>
<td>.810</td>
<td>.117</td>
<td>.003</td>
<td>-.085</td>
<td>.055</td>
<td>-.020</td>
<td>-.076</td>
<td>.069</td>
</tr>
<tr>
<td>Self-Management</td>
<td>.109</td>
<td>.580</td>
<td>.122</td>
<td>.084</td>
<td>.099</td>
<td>-.203</td>
<td>-.173</td>
<td>.070</td>
</tr>
<tr>
<td>Verbal</td>
<td>-.119</td>
<td>.027</td>
<td>.072</td>
<td>.424</td>
<td>.574</td>
<td>-.014</td>
<td>.093</td>
<td>-.036</td>
</tr>
<tr>
<td>Clerical</td>
<td>.184</td>
<td>.558</td>
<td>.076</td>
<td>-.094</td>
<td>-.148</td>
<td>.101</td>
<td>.036</td>
<td>-.135</td>
</tr>
<tr>
<td>Stress Resistance</td>
<td>.214</td>
<td>.357</td>
<td>.157</td>
<td>.084</td>
<td>.254</td>
<td>.200</td>
<td>-.138</td>
<td>-.077</td>
</tr>
<tr>
<td>Math</td>
<td>.500</td>
<td>.357</td>
<td>.367</td>
<td>-.176</td>
<td>-.344</td>
<td>.049</td>
<td>.049</td>
<td>-.188</td>
</tr>
<tr>
<td>Spatial</td>
<td>.687</td>
<td>.007</td>
<td>.113</td>
<td>-.081</td>
<td>.044</td>
<td>.037</td>
<td>-.164</td>
<td>.343</td>
</tr>
<tr>
<td>Science</td>
<td>.782</td>
<td>.075</td>
<td>.160</td>
<td>.128</td>
<td>.069</td>
<td>-.205</td>
<td>-.048</td>
<td>-.074</td>
</tr>
<tr>
<td>Self-Ratings of Ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math-Spatial</td>
<td>.715</td>
<td>.349</td>
<td>.268</td>
<td>-.301</td>
<td>-.121</td>
<td>.100</td>
<td>-.010</td>
<td>-.093</td>
</tr>
<tr>
<td>Verbal</td>
<td>-.089</td>
<td>.591</td>
<td>-.263</td>
<td>-.131</td>
<td>.177</td>
<td>.093</td>
<td>-.145</td>
<td>.008</td>
</tr>
<tr>
<td>Self-Regulation</td>
<td>-.093</td>
<td>.032</td>
<td>.052</td>
<td>.185</td>
<td>.927</td>
<td>.192</td>
<td>.143</td>
<td>-.118</td>
</tr>
<tr>
<td>Ability Measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>.285</td>
<td>.033</td>
<td>.530</td>
<td>-.042</td>
<td>-.007</td>
<td>-.003</td>
<td>.002</td>
<td>.664</td>
</tr>
<tr>
<td>Verbal</td>
<td>.084</td>
<td>-.080</td>
<td>.528</td>
<td>.537</td>
<td>.402</td>
<td>-.269</td>
<td>.121</td>
<td>-.044</td>
</tr>
<tr>
<td>PS-Complex</td>
<td>.291</td>
<td>.045</td>
<td>.809</td>
<td>.069</td>
<td>.004</td>
<td>.070</td>
<td>-.103</td>
<td>.025</td>
</tr>
<tr>
<td>PS-Simple</td>
<td>-.003</td>
<td>-.094</td>
<td>.713</td>
<td>.109</td>
<td>.127</td>
<td>.057</td>
<td>-.127</td>
<td>.046</td>
</tr>
<tr>
<td>Math</td>
<td>.391</td>
<td>.219</td>
<td>.699</td>
<td>-.028</td>
<td>-.077</td>
<td>-.206</td>
<td>-.009</td>
<td>.203</td>
</tr>
<tr>
<td>Motivational Skills</td>
<td>.261</td>
<td>.225</td>
<td>.141</td>
<td>.047</td>
<td>.298</td>
<td>.225</td>
<td>-.225</td>
<td>.008</td>
</tr>
</tbody>
</table>

Note. Salient loadings are indicated in boldface. TIE = Typical Intellectual Engagement; UNIACT = Unisex edition of the ACT Interest Inventory; PS = Perceptual Speed.

domains, but especially among self-concept, vocational interests, and some measures of personality. However, Neuroticism and Anxiety, though highly intercorrelated with one another, show little overlap with other trait domains under consideration here.

One finding from these data that has some controversial history in the self-concept literature (e.g., Marsh, 1986, 1990; Skaalvik & Rankin, 1990) is the negative loadings of math self-concept and math/spatial self-ratings of ability on factors that have high salient positive loadings of verbal self-
concept and verbal self-ratings of ability. Negative correlations were found between math self-concept and math–spatial self-ratings of ability on the one hand and verbal self-concept and verbal self-ratings of ability on the other hand. That is, in this sample of college students, there appears to be some “ipsatizing” of the math–verbal continuum (e.g., if individuals believe that if they are high in verbal abilities, then they are low on math abilities). This may reflect trainees’ preference for one type of material or another or a more pervasive implicit sense of compensatory abilities. Raw correlations show Math Ability and Math Self-Concept to be similarly associated with one another \((r = .55)\), as are Verbal Ability and Verbal Self-Concept \((r = .54)\), even though Math Self-Concept correlates with Verbal Self-Concept \((r = .28)\), and Math Ability correlates with Verbal Ability \((r = .38)\).

Postscript: PS-Complex Ability

Given the dominant influence of the PS-Complex ability measure in predicting performance in the current study and in previous studies for complex skill performance (e.g., Ackerman & Kanfer, 1993b; Cobb & Mathews, 1972; Guilford & Lacey, 1947), we performed a set of follow-up analyses to evaluate communality of PS-Complex with nonability distal measures. The primary rationale for this analysis was the observation (from a set of performance–stress studies) that the often frustrating, iterative strategy necessary for successful performance of the Dial Reading and Directional Headings tests may yield a performance-based assessment of individual differences in susceptibility to emotion control deficits. We first subjected PS-complex (made up of the Dial Reading and Directional Headings tests) to cross-correlation analyses with nonability predictors of performance in TRACON. Similar to correlations we have found in an earlier study (Ackerman & Kanfer, 1994), PS-Complex correlated significantly with several nonability measures as follows: Anxiety \((r = -.21)\), TIE \((r = .22)\), Realistic Interests \((r = .35)\), and Spatial–Mechanical–Science Self-Concept \((r = .34)\). For proximal and concomitant measures, PS-complex also correlated significantly with task self-efficacy (performance) across practice \((r_s = .37, .33, .38, .39, .36, .31)\) for Sessions 1–6 respectively. Of considerable importance were the correlations between PS-Complex and initial negative motivation and positive motivation thoughts. Consistent with our interpretation of the emotion-control influences on performance in the Dial Reading and Directional Headings tests, PS-Complex correlated negatively with negative motivation scores \((r = -.31)\), indicating that low scores on PS-Complex were associated with greater reporting of negative motivational thoughts prior to task practice. In addition, PS-Complex showed essentially a zero correlation with positive motivational thoughts prior to Session 1 \((r = .07)\), even though negative and positive motivation thoughts were highly negatively correlated \((r = -.58)\).

Also, a further result that was consistent with our speculation about the tests, PS-Complex correlated significantly with a self-report measure of Stress-Resistance Self-Concept \((r = .28)\). PS-Complex was the only objective ability test composite that significantly correlated with this nonability measure.

Finally, we subjected PS-Complex to a gender analysis. Even though the constituent tests have some surface-level spatial content (which would lead to an expectation of higher mean scores for men), only a marginally significant difference was found for gender; \(M_{Women} = -.18, M_{Men} = .22, t(91) = 1.96, p = .05\).

When these results are considered in light of the previous literature on validity of the Dial Reading and Directional Headings tests for predicting performance, it seems clear that a class of objective tests may provide insight into ability–nonability relations and may also provide a source for both construct and criterion-related validity above and beyond so-called \(g\) measures. Such findings also may illuminate why some PS ability measures have high predictive validity for intermediate stages of skill acquisition (e.g., Ackerman, 1988, 1990) and for the individual-differences construct identified by Ackerman and Woltz (1994) as “propensity-to-learn.” (See Ackerman & Goff, 1995, for a review of ability–nonability relations.)

Discussion

Individual differences in performance on a complex skill acquisition task are well-predicted by a variety of distal cognitive and noncognitive trait
measures, including ability, interests, motivational skills, self-concept, and self-ratings of ability. In addition, proximal measures of negative and positive motivational thoughts and self-efficacy were also substantially related to task performance. However, substantial overlap in variance was found among the distal variables, proximal variables, and between distal and proximal variables. By simultaneous examination of these variables in predicting task performance, early and late in practice, we demonstrated that much of the valid criterion-related variance among nonability predictors was variance that was shared with cognitive ability variables. Communalities among the cognitive and noncognitive determinants of performance revealed by MDS and factor analysis suggest that aptitude complexes for skill acquisition can be identified and, in turn, used in an integrated selection program (for a discussion of aptitude complexes, see Snow, 1989). In addition, the demonstrated communalities may provide a further basis for building models of learner-task relations that better represent the interrelations of cognitive and noncognitive traits. Models that take these cognitive and noncognitive determinants into consideration may ultimately allow for a joint prediction of the effort allocated to a task and strategic decisions made by trainees (e.g., choice behaviors that are guided by self-concept or self-efficacy).

The substantial effort in this experiment devoted to pretask ability testing and assessment of individual differences in a variety of nonability predictors of performance provided for a wider assessment of the predictor construct space and the joint assessment of performance on a complex skill-learning task than has been possible in previous studies. In many ways, the TRACON task is representative of the kind of real-world tasks for which performance under time stress is observed. The task is continuous (over a 30-min watch), can involve unpredictable events, and has substantial information processing stress (by virtue of aircraft pilots giving repeated requests for permissions over the headset and when separation conflicts are in progress and warning alerts are broadcast visually and auditorially). The only salient aspects of the task that are not characteristic of the real-world is that no actual loss of human life occurs when mistakes are made, and no direct threat is imposed on the individual operator. Nonetheless, we have observed behaviors in the laboratory that are consistent with the inference that the task is stressful. Incidences of trainees who become upset in conjunction with poor performance—who curse or become agitated with the equipment (e.g., pounding the keyboard or trackball)—have been observed over the course of our experiments with TRACON. The observed negative motivation thought occurrences between task sessions are consistent with this inference as well. Some trainees report worry and apprehension prior to the TRACON sessions and report negative affective reactions at the conclusion of TRACON sessions. Although there is controversy in the field about whether the job of air traffic controller is of greater stress than other jobs (e.g., see Hopkin, 1988, for a review), there is virtually no doubt that the task is stressful during the skill acquisition phase for many trainees.

Conclusions

This study indicated that many ability and nonability predictors provide significant correlations with individual differences in TRACON performance over practice. The nonability predictors range from distal vocational interests to academic and ability self-concept, motivational skills to proximal reports of negative and positive motivational thoughts and self-efficacy expectations. However, when objective ability measures are first entered into a multiple regression equation for predicting task performance, little incremental validity is found for the nonability predictors. In this study, only incidents of proximal negative and positive motivational thoughts consistently added significant incremental prediction to the equation. That is, common variance among most of the nonability predictors, ability measures, and TRACON performance is accounted for by several objective ability composite measures.

Most prominent among these ability predictors were the two tests identified as PS-Complex, namely the Dial Reading and the Directional Headings tests. However, it is important to note that such variables also share considerable common variance with the important nonability predictors of performance. These results support the inference that both ability and nonability influences of individual differences in performance under stressful conditions may be well accounted for by a set of objective tests with characteristics similar to the
Dial Reading and Directional Headings tests. Although there is a literature concerning these tests, investigators have not examined the nonability determinants of performance on these tests, even though these tests have demonstrated the highest validity in large batteries of ability tests for predicting job performance under stress. Our results strongly suggest that future investigations should be conducted to broaden the construct space underlying these measures, for example, by developing a battery of tests that have similar information-processing characteristics. With a larger battery of these tests, it may be possible to account independently for the ability and nonability determinants of performance on such measures. Such a result could have implications for new assessment instruments for selection of applicants for jobs that involve substantial stress exposure.

An extensive set of analyses of the predictor space illustrated many sources of communality across domains that are often implicitly considered in the literature to be orthogonal. Specific associations between personality and interest measures, between self-concept and ability measures, and between abilities and both interests and self-concept are particularly illuminating, especially for those researchers who are interested in predictions of individual differences in task performance. That is, in the absence of ability predictors, many measures of self-concept, vocational interests, and self-efficacy estimates show significant correlations with complex task performance across task practice. However, when a broad array of ability measures is entered into a regression equation first, few of these nonability measures provide incremental predictive validity for performance. Of course, an alternate viewpoint would suggest that these nonability predictors can serve as surrogates for objective ability measures. For example, Math, Spatial, and Verbal Self-Concept measures correlated with the objective ability composites ($r = .55, .51, \text{ and } .55$, respectively).

Traditional approaches to prediction tend to emphasize ability or nonability domains. The pattern of results obtained in this study suggests that these traditional perspectives may obscure important relations among predictors and performance criteria. For example, path analyses indicate the significant role of both ability and nonability influences for predicting performance. From a traditional ability perspective, the results provide confirmation of ability influences on performance. From a traditional nonability perspective, the results provide support for the notion of significant nonability influences on performance (by means of proximal measures that stem from integration of contemporary motivation-based theories of performance). But what is typically neglected by both perspectives is the critical linkage between individual differences in abilities and individual differences in related interests, self-concepts, and proximal motivational measures, such as self-efficacy. Although theorists have long posited such a linkage (Cattell, 1946; Wechsler, 1950), empirical evidence that supports this type of ability–nonability linkage has been sparse (see Ackerman & Goff, 1995, for a review).

The link between ability and nonability predictors has important implications for broader conceptual issues and for prediction of performance as follows:

1. For prediction of performance, individual differences in abilities obviously play a key role. Ability measures accounted for roughly 30% to 45% of the variance in complex task performance over practice. However, an individual's perceptions of these abilities as they relate to diverse domains of behavior may importantly mediate the ability–performance relations in unforeseen ways, particularly when successful performance over time involves emotion control and persistence. That is, objective ability measures are moderately-to-highly correlated with self-concept and self-estimates of ability, the self-concept measures have a direct influence on negative and positive motivational thoughts and on task self-efficacy. How trainees view their abilities may be important in determining how they confront a novel task, but the objective measures of ability predict better how well they actually perform the task.

2. Controversy over the association among measures of major personality dimensions and performance has focused on matching personality measures (grounded in theories of personality structure) to the performance (criterion) space. Results of this study suggest another fruitful avenue for resolution of these issues pertains to a broadening of the predictor domain to include measures and concepts from vocational psychology
COGNITIVE AND NONCOGNITIVE DETERMINANTS

and self-concept theories. In the integrated perspective presented here, self-concept and interest measures represent constructs with important links to broad personality dimensions, as well as performance criteria.

3. In applied settings, thorough assessment of individual differences in cognitive abilities is often precluded for a variety of reasons. Individual differences in nonability predictors may provide a useful partial proxy for such assessments when ability data are unavailable. However, as shown in the present results, when performance is complex and the demands of the task change over practice, such proxy measures may introduce systematic, non-performance-related variance into the equation. Indeed, we found gender differences in task self-efficacy, even after variance attributable to ability, self-concept, interests, and the like had been removed. It remains to be seen what determines these residual differences in self-efficacy and whether such differences in turn influence choice decisions on how to engage the task assignment. It will be important in future investigations to discover the sources of discrepancies between individual differences in task performance and task self-efficacy.

4. Measures typically considered to be ability predictors (such as the Dial Reading and Directional Headings tests) may indeed capture substantial variance in the nonability domain. Questions of the relative effectiveness of these measures for determining nonability variance, compared with self-report nonability measures depend, of course, on the construct under consideration. However, these initial findings provide a potentially important direction for future research on performance-based assessment of nonability constructs. In an applied setting such as job selection, concerns about demand characteristics (e.g., “faking good”) on nonability measures might be circumvented by administering objective ability tests that tap individual differences in nonability constructs, given that it is generally impossible to fake good on an ability test. The high validity demonstrated by the PS-Complex measure suggested that this particular trait may represent an especially useful addition to more traditional batteries of cognitive abilities in predicting individual differences in performance on complex skill tasks.

References


Ackerman, P. L., & Goff, M. (1995). Intelligence, personality, and interests: Historical review and a model for adult intellect. Unpublished manuscript under review.


Adams, J. A. (1987). Historical review and appraisal of
research on the learning, retention, and transfer of human motor skills. Psychological Bulletin, 101, 41–74.


Mitchell, T. R., Hopper, H., Daniels, D., George-Falvey, J., & James, L. R. (1994). Predicting self-efficacy and


**Appendix**

**TRACON Simulation Task**

The task used for this research is a uniquely modified professional version (V1.52) of TRACON simulation software developed by Wesson International. Versions of this program are in use in several locations in the United States (including the FAA, National Aeronautics and Space Administration and Department of Defense and several colleges and university airway sciences programs) for training of air traffic controllers. Modifications for the current instantiation of the program have allowed for the collection of a variety of data described in more detail below.

Descriptions of the TRACON simulations are provided in Ackerman (1992) and Ackerman and Kanfer (1993b). The following discussion abstracts that material with deviations specifically noted. The task requires that trainees learn a set of rules for air traffic control, including reading flight strips, declarative knowledge about radar beacons, airport locations, airport tower handoff procedures, en route center handoff procedures, plane separation rules and procedures, monitoring strategies, and strategies for sequencing planes for maximum efficient and safe sector traversal. In addition, trainees are required to acquire human–computer interface skills that include issuing trackball-based commands, menu retrieval, keyboard operations, and integration between visual and auditory information channels. However, the simulator task represents a substantial reduction of rules and operational demands in comparison to the real-world job of air traffic controllers.
Display

TRACON presents the trainee with a simulated color radar screen depicting a region of airspace, very high frequency omnidirectional range stations (VOR), airports, sector boundaries, and range rings. Planes are identified by icon on the radar scope, with a data tag that indicates plane identification and altitude information. In addition, two sets of flight strips are presented at the right side of the display, one “pending” and the other “active.” Each flight strip contains information about a particular flight, including identification information, plane type, requested speed and altitude, and sector entry and exit destination information (see Figure A1). At the bottom of the screen is a communications scope.

Figure A1. Static copy of TRACON screen. There are three major components to the display. The right-hand side of the screen shows pending (not under control) and active (under control) flight strips. Each flight strip lists (a) plane identifier, (b) plane type, (c) requested speed, (d) requested altitude, (e) radar fix of sector entry, (f) radar fix of sector exit (including tower or center). The lower part of the screen shows a communications box that gives a printout of the current (and last few) commands issued by the trainee and the responses from pilots or other controllers. The main part of the screen shows a radar representation of the Chicago sector. Planes are represented by a plane icon and a data tag, which gives the identifier, the altitude, and an indication of current changes in altitude. The sector is bounded by the irregular dotted polygon describing a perimeter. Radar fixes are shown as small (+) figures on the radar screen. Airports are shown with approach cones and a circle indicating the facility proper. A continuous radar sweep is shown (updating at 12 o’clock every 5 s). Range rings are also displayed indicating 5-mile (8.0467-km) distances.
“communications box,” which shows commands issued to planes and pilots responses, along with the controller’s score for the current simulation. When planes are about to enter the trainee’s sector (at a boundary or on the runway of an airport), this information is announced over the headset. No flight is allowed to cross the sector boundary or take off from an airport without explicit authorization by the trainee.

Task Controls and Knowledge of Results

Trainees interacted with the TRACON simulation in several ways. A trackball was used for the majority of input activities, although the keyboard was also used alone or in conjunction with the trackball. (The trackball represents a change in input device from our study where a mouse was used). For each plane command, a menu of command choices is displayed on the screen.

Knowledge of results is provided visually by text in the communications box and aurally with a read-back by the pilot or other controller using digitized speech. In addition, planes follow as nearly as possible the commands issued by the trainee. Commands such as turn, change altitude, and change speed are processed by the computer and carried out in accordance with the limitations imposed by each aircraft type.

When errors occur (e.g., separation conflicts, near misses, crashes, missed approaches, or handoff errors), additional information is presented to the trainee. In each of these cases, an alert circle around the plane or planes in question is presented on the screen, and a series of tones are presented over the headset.

Trial Description

Trials for the task have been created and pretested to be roughly equivalent in difficulty. Each trial contained planes that were divided into three basic categories (overflights, departures, and arrivals). Overflights are planes that enter and exit the trainee’s airspace at cruising altitudes. Trainees are required to acknowledge these airplanes as they approach a boundary VOR fix, monitor progress through the sector, and handoff to a “center” controller. Departures are planes that originate at one of the four airports, climb to a cruising altitude, and are handed off to a center controller. Trainees are required to release departures from airports, evaluate and remediate potential conflicts as the planes climbed to a cruising altitude and turn to intercept their intended flight paths, and then handoff planes to the appropriate center controller. Arrivals enter the trainee’s airspace from one of the boundary VOR fixes and are to be landed at a designated airport. Trainees are required to direct arrivals onto an appropriate heading and altitude to provide an acceptable handoff to the appropriate tower controller before these planes land. For all flights, the trainee is required to maintain legal separation, that is at least 1,000 ft. (304.8 m) in altitude or 3 miles (5.56 km) horizontally.

Each trial lasted 30 min and consisted of 16 overflights and departures (with roughly equal frequency) and 12 arrivals. The planes request entry to the airspace at irregular intervals constrained so as to require the trainee to be continuously occupied with at least one active target. The trials were designed so that perfect performance (handling all 28 planes successfully) was just beyond the skill level achieved by subject matter experts.

A successful “handle” of a flight is the appropriate accomplishment of the respective flight plan. That is, for a departure or an overflight, the accomplishment is a successful handoff to the appropriate center controller. For a landing, the accomplishment is the successful landing of the airplane.

Performance Measurement

As with a previous investigation (Ackerman, 1992), overall performance was computed as the sum of all flights accepted into the sector that had a final disposition within the simulation time minus any planes that were incorrectly disposed of (e.g., crashes, not-handed-off, or vectored off the radar screen). This measure is concordant with measures derived from the examination of the criterion space for FAA air traffic controller simulation research (e.g., see Buckley, Debaryshe, Hitchner, & Kohn, 1983) and has acceptable reliability.

Received March 1, 1995
Revision received May 22, 1995
Accepted May 25, 1995