Part III

Theoretical Frameworks for Specific Domains Not Included in ALL and for Which Assessment Tools Remains to be Developped

This part of the publication includes three chapters. Chapters 7 and 8 provide assessment frameworks for two skill domains where development failed to yield approaches to measurement that were sufficiently robust to meet the criteria set out for inclusion in the international comparative assessment. These chapters also set out what was learned during the process of development and validation. Chapter 7 presents the framework for teamwork and chapter 8 presents the framework for practical cognition. Chapter 9 provides the assessment framework developed by the Educational Testing Service (ETS) for measuring information and communication technology literacy. Although it was not available soon enough to inform the design of the ALL study, pilot testing has revealed that the framework is viable and the approach to measurement yields robust estimates that would meet the demanding criteria set for ALL.

Chapter 7

The ALL Teamwork Framework

David P. Baker *American Institutes for Research, U.S.*
Lisa Horvath *George Washington University, U.S.*
Michael Campion, *Purdue University, U.S.*
Lynn Offermann *George Washington University, U.S.*
Eduardo Salas *Orlando, FL, U.S.*

Table of Contents

Table of Contents

Summary

Governments, businesses, and community groups are increasingly relying on work teams to streamline processes, enhance participation, and improve performance. Teamwork is of worldwide importance; individuals who wish to participate fully in community and professional life must increasingly possess the skills necessary to work in teams.

Although teams are diverse and can take on many forms, all teams are defined by four characteristics. They have two or more individuals; they share a common goal(s); they are task-interdependent; and they have a desired productive outcome(s). These characteristics serve as the basis for developing a working definition of a "team," a definition that the ALL can use to provide insight regarding the prevalence and the expression of teamwork skills across various cultures.

The Teamwork scale of ALL seeks to assess the core skills associated with teamwork. To this end, three primary skills required for effective teamwork—Group Decision Making/Planning, Adaptability/Flexibility, and Interpersonal Relations—are proposed, each represented by distinct behavioral manifestations. Group Decision Making/Planning refers to the ability to identify problems and gather, evaluate, share and link information. Adaptability/Flexibility implies using a variety of task-relevant strategies, providing assistance, adjusting to task reallocation and accepting feedback. Interpersonal Relations reflects supporting team decisions, sharing work, helping others, and seeking mutually agreeable solutions. Communication skills—including providing complete and concise information, listening effectively, and asking questions—underlie the other three skills and serve as a bridge among them. In addition, two other factors play key roles in teamwork: attitudes toward teamwork and past experience with teams.

By definition, teamwork skills can only be observed directly in a teamwork setting. However, because direct observation is not consistent with ALL methodology, respondent teamwork skills should be assessed indirectly. Specifically, knowledge of teamwork skills, attitudes towards teamwork, and past experience in teams should be measured, and links should be drawn between these constructs and team performance.

Finally, teamwork, more than other life skills, is likely to be affected by culture. Although the team skills described in this framework are assumed to define teamwork generally, the behavioral manifestation of these skills is likely to vary across cultures. Respondent performance should be interpreted relative to the effective teamwork behaviors defined for a given country, thereby providing information regarding national attitudes toward teamwork and regarding the degree to which behavioral expressions of teamwork skills vary across nations. This information can be used by employers and educators alike, to assess and improve teamwork in a nation's workforce and general population.

1. Introduction

This document presents a framework for assessing Teamwork as part of the Adult Literacy and Life skills survey (ALL). The framework was developed from the literature on teams and what is currently known about teamwork. Overall, the framework serves three purposes. First, it bounds the problem domain by clearly specifying the critical components of teamwork to assess. Our goal here is to target the most fundamental aspects of teamwork. Second, the framework drives our approach to measurement. Strategies that are most effective for assessing team knowledge, skills, and attitudes should be identified and selected. Finally, based on the measurement strategies identified, the framework serves as the template for item development. Items should be developed to target key aspects of teamwork that are specified in this framework.

The framework is divided into five sections. The first presents a detailed discussion of the literature on teams and what is currently known about the knowledge, skills, and attitudes required for effective team performance. Here, we present our definition of a team and clearly delineate the core facets of teamwork. Rather than including all variables, we present the core dimensions that characterize what teams do. These dimensions are assumed to be central to all teams, regardless of culture.

In the second section, we draw on the results of our literature review to build a model of teamwork. The purpose of this model is to identify key areas for measurement in ALL. In particular, we propose that the teamwork measure should assess what team members bring to a team (e.g., attitudes, past experience, etc.) and what team members do in a team (e.g., interact, coordinate, etc.).

Our proposed framework is intended to target the most fundamental aspects of teamwork and recognizes that the primary goal of the teamwork measure is not to assess differences in culture. Nonetheless, given the interpersonal nature of teamwork, we anticipate cultural differences. The third section of this framework addresses this issue. In particular, relevant cultural research is reviewed, and the implications of these studies are discussed in light of our objective (i.e., measuring team knowledge, skills, and attitudes internationally).

Once the key facets of teamwork are identified and the possible effects of culture are discussed, the fourth section of this framework presents specific strategies for measuring teamwork. We first present the theoretical and practical assumptions that guide our approach. Next, we describe each proposed teamwork measure with respect to the measurement approach employed, the process by which items were developed, and the procedures for scoring.

Finally, the fifth section of this framework briefly discusses social and economic indicators that may affect teamwork. Here, variables are proposed for the respondent background questionnaire. Information on these variables, which are expected to moderate participant responses on the teamwork measure, should also provide insights into the determinants of teamwork in different nations.

1.1 Why measure teamwork?

Organizations (both work and non-work) are increasingly using teams to streamline processes, enhance participation, and improve quality (Cohen and Bailey, 1997). Hence, teams are becoming the primary building block of most organizations (Brooks, 1993; McGrath, 1997). In fact, a recent study by Gordon (1992) found that 82% of U.S. companies with 100 or more employees utilize some form of teams. Teams are found in such diverse fields as education, religion, science, manufacturing, and consulting.

Because teams span both private and public life, individuals must be able to work and perform in a team context to function effectively in today's society. Both the Secretary's Commission on Achieving Necessary Skills (SCANS; U.S. Department of Labor, 1991, 1992a, 1992b) and the Conference Board of Canada Employability Skills Profile (1993) cite the importance of interpersonal skills (or teamwork) in work and everyday life.

Due to its prevalence in society, teamwork has been identified as an important life skill. Consistent with the goals of ALL, the teamwork measure should provide information as to how teamwork skills are distributed in the adult population internationally. Information on the nature of teamwork skills associated with a particular nation and the social and economic factors that influence the development of teamwork skills are of particular interest. This information should prove valuable to employers and educators who wish to improve teamwork in the workforce and elsewhere.

1.2 Challenges of the project

Although there is little doubt that teamwork is an important life skill, the measurement of teamwork in ALL presents specific challenges. First, ALL will be the initial attempt to provide a large-scale international assessment of teamwork skills. As a result, a limited number of methods and approaches exist as precedents. Past international assessments have focused on adult literacy (i.e., see IALS) as opposed to interpersonal skills like teamwork. Therefore, we expect to learn a great deal about teamwork skills and their distribution in the adult population across nations.

Second, unlike other life skills measured by ALL, teamwork will likely be affected by culture. Although we believe that a certain set of core skills defines teamwork across all cultures, the way in which these skills are manifested within a team is likely to vary. Therefore, we will not attempt to develop an invariant set of items to be translated for use in each nation. Rather, items should be modified as necessary to take known cultural differences into account.

Finally, and perhaps most challenging, is the fact that most methods of assessing teamwork skills require direct observation of team performance (D. Baker and Salas, 1992; 1997; Brannick, Prince and Salas, 1997; Ilgen, 1999). Typically, team members are placed in a scenario. Experts observe team behaviors and provide performance ratings on specific teamwork skills. This measurement approach differs substantially from the approach to be used in ALL. Here, the teamwork measure will be a short paper-based instrument; no opportunity for directly observing the teamwork skills of the respondent will be available. Thus, respondents' teamwork skills will have to be assessed indirectly instead of directly. For ALL, we propose to measure the knowledge of teamwork skills, attitude towards teamwork, and past experience in teams and then attempt to draw links between these variables and team performance.

With these challenges in mind, we turn to a discussion of the key components of teamwork underlying our framework. Because these domains will drive development of the teamwork measures for ALL, we draw heavily from the literature on teams and on what is currently known about teamwork.

2. Teams and teamwork

2.1 What is a team?

Although a widespread consensus acknowledges the prevalence of teams in society, the research literature reflects only marginal agreement concerning the definitional components of teams. The variance in definitions is due in part to the diversity of team types. Teams carry a variety of purposes (e.g., learning, producing a product, solving problems, gaining acceptance), forms (e.g., virtual, co-located), and sizes and longevity (e.g., adhoc, long term) (Cohen and Bailey, 1997).

In an attempt to extract key features of teams and develop a working definition of teams for ALL, we reviewed several often-cited definitions (Dyer, 1984; Guzzo and Shea, 1992; Mohrman, Cohen, and Mohrman, 1995; Salas, Dickinson, Converse and Tannenbaum, 1992). This process produced four common characteristics of a "team."

- Two or more individuals
- A shared or common goal(s)
- Task interdependency
- A desired productive outcome(s)

These characteristics serve as the basis for developing our working definition of a "team." A clear definition of a team is essential because it provides measurement boundaries and clearly distinguishes teams from small groups, which do not necessarily connote interdependence. (A team is also a "small group," but a small group may or may not be a team.) Our definition of a team is as follows:

> *A team consists of two or more individuals who must interact to achieve one or more common goals that are directed toward the accomplishment of a productive outcome(s).*

In addition, the definition and core characteristics provide preliminary insight into the nature of teamwork and its key facets. For example, the characteristics of task interdependency and shared goals imply that team members must collectively decide on team goals (team decision making) and work cooperatively (coordination) to achieve these goals.

2.2 What is teamwork?

Teamwork has traditionally been described in terms of classical systems theory in which team inputs, team processes, and team outputs are arrayed over time. Here, team inputs include the characteristics of the task to be performed, the elements of the context in which teamwork occurs, and the attitudes team members bring to a team situation. Team process includes the interaction and coordination among members required for performing team tasks and achieving specific goals. Team outputs consist of the products that result from team performance (Hackman, 1987; Ilgen, 1999; McGrath, 1984). With regard to teamwork, the process phase is the defining point at which teamwork occurs; it is during this phase that team members interact and work together to produce team outputs.

Numerous theories have been proposed and extensive research has been conducted on the nature of team process (i.e., teamwork). Historically, this literature has sought to identify generic teamwork skills that are associated with most teams. More recently, the

focus has shifted towards researchers identifying the specific *competency requirements* of team members (Cannon-Bowers, Tannenbaum, Salas, and Volpe, 1995; O'Neil, Chung, and Brown, 1997; Stevens and Campion, 1994). The term *competency* has a variety of meanings. However, it is generally used to denote the qualities needed by a jobholder (Boyatzis, 1982)[1]. Specifically, Parry (1998) defined the term "competencies" as a cluster of related knowledge, skills, and attitudes that affects a major part of one's job (i.e., one or more key roles or responsibilities); is correlated with performance on the job; can be measured against well-accepted standards; and can be improved through training and development.

Regarding teamwork, team competencies are the qualities needed by team members. Cannon-Bowers et al. (1995) identified three types of competencies that are central for effective teamwork: (1) team knowledge competencies, (2) team skill competencies, and (3) team attitude competencies.

Team Knowledge Competencies. Team knowledge competencies are defined by Cannon-Bowers et al. (1995) as the principles and concepts that underlie a team's effective task performance. To function effectively in a team, team members must know what team skills are required, when particular team behaviors are appropriate, and how these skills should be utilized in a team setting. In addition, team members should know the team's mission and goals and be aware of each other's roles and responsibilities in achieving those goals. Such knowledge enables team members to form appropriate strategies for interaction, to coordinate with other team members, and to achieve maximum team performance.

Team Skill Competencies. Team skill competencies, which have received considerable research attention, are defined as a learned capacity to interact with other team members at some minimal proficiency level (Cannon-Bowers et al., 1995). However, Cannon-Bowers et al. has reported that the literature on team skills is confusing and contradictory, as well as plagued with inconsistencies in terms of both skill labels and definitions. Across studies, different labels are used to refer to the same teamwork skills or the same labels are used to refer to different skills. In an attempt to resolve these inconsistencies, Cannon-Bowers et al., found that 130 skill labels could be sorted into eight major teamwork skill categories: adaptability, situation awareness, performance monitoring/ feedback, leadership, interpersonal relations, coordination, communication, and decision making. Numerous investigations have shown that these skills are directly related to team performance (see for example, Morgan, Glickman, Woodward, Blaiwes, and Salas, 1986; Oser, McCallum, Salas, and Morgan, 1992; Salas, Bowers, and Cannon-Bowers, 1995; Salas, Fowlkes, Stout, Milanovich, and Prince, 1999).

Team Attitude Competencies. Team attitude competencies are defined as an internal state that influences a team member's choices or decisions to act in a particular way (Cannon-Bowers et al., 1995; Dick and Carey, 1990). Attitudes toward teamwork can have a significant effect on how teamwork skills are actually put into practice. Positive attitudes toward teamwork and mutual trust among team members are examples of critical attitudes related to team process (Gregorich, Helmreich and Wilhelm, 1990; Ruffell-Smith, 1979; Helmreich, Fushee, Benson, and Russini, 1986). For example, Vaziri, Lee, and Krieger (1988) found that higher levels of mutual trust among team members led to a more harmonious and productive team environment. Finally, an attraction to being part of a team (i.e., collective orientation) is critical (Eby and Dobbins, 1997). Driskell and Salas (1992) reported that collectively-oriented individuals performed significantly better than did individually-oriented team members because

1. Boyatzis (1982), in his seminal work on competencies, defines a job competency as "an underlying characteristic of a person, which results in effective or superior performance in a job."

collectively-oriented individuals tended to take advantage of the benefits offered by teamwork. Furthermore, collectively-oriented individuals had the capacity to take other team members' behavior into account and believed that a team approach was superior to an individual one.

Refining the work of Cannon-Bowers et al. (1995), Cannon-Bowers and Salas (1997) delineated three types of team knowledge, skills, and attitude competencies. First, "*individual* competencies" are defined as the knowledge, skills, and attitudes required on the part of individual team members to perform position requirements. These competencies enable team members to perform tasks that are specifically assigned to them. For example, an individual in a marketing team assigned to purchase newspaper-advertising needs to possess specific knowledge and skills to successfully perform this task. Second, "*team* competencies held at the *individual level*" are defined as the knowledge, skills, and attitudes that are generic with respect to a team and its tasks. Essentially, these competencies are transportable to different teams and different team settings. For example, knowledge about teamwork skills and behaviors; skill in communication, team decision making, and interpersonal relations; positive attitudes toward teamwork, and a collective orientation enable team members to function effectively across a wide variety of teams. Finally, "*team* competencies held at the *team level*" are defined as the knowledge, skills, and attitudes that are specific to a particular team and task. Unlike team competencies at the individual level, these competencies are not transportable. They only have meaning within the team. For example, knowledge of teammate roles and responsibilities and specific teammate characteristics are only useful within a specific team context.

Given that the primary goal of ALL is to assess teamwork in the adult international population, teamwork measures will assess "*team* competencies held at the *individual level*." By definition, these competencies are of great interest to policymakers and educators because they enable individuals to function effectively in a wide variety of teams and a wide variety of team settings.

2.3 Core team skills, knowledge, and attitudes

A comprehensive review of teamwork models and research was conducted (e.g., Carnevale, Gainer and Meltzer, 1990; Commission on the Skills of the American Workforce, 1990) to identify core team knowledge, skills, and attitude competencies held at the individual level. From that broad review, we selected the most comprehensive and current team competency models (Cannon-Bowers et al., 1995; O'Neil et al., 1997; Stevens and Campion, 1994a) and used these models to identify core team competencies to measure in ALL. Competencies were selected based upon the following criteria: (1) the competencies were held at the individual level; (2) at least two of the three models delineated the competency (in some form); and (3) empirical research supported a positive relationship between the competency and performance.

Core Team Skills. Team skill competencies are discussed first because they represent the manifest, individual-level behaviors that the ALL measure is designed to assess. Four competencies were identified as "core" team skills competencies: communication, interpersonal relations (which includes cooperation and dealing with conflict), group decision making/planning, and adaptability/flexibility. Team leadership, an often-cited skill competency (see for example, Cannon-Bowers et al., 1995), was not included because our current focus is on the ability to *work* in a team, not to *lead* one. Each core team skill is defined below, along with behavioral examples that typify the skill's expression. Although this core is assumed to reflect teamwork in most cultures, it should be noted that the behavioral exemplars presented here were derived from research

conducted on teams in the U.S. (Cannon-Bowers et al., 1995; O'Neil et al., 1997; Stevens and Campion, 1994a). As such, they may or may not be consistent with the *expression* of the same core skills in other cultures. Thus, the cross-cultural generalizability of behaviors that manifest core team skills in the U.S. remains an empirical question that the ALL will address. However, should cultures to which these behaviors do not generalize be included in the ALL, other behaviors are expected to express the same core team skill competencies systematically.

Communication is defined as establishing effective communication between self and others; it involves the exchange of clear and accurate information and the ability to clarify or acknowledge the receipt of information.

Strong communication skills are demonstrated by team members who

- Provide clear and accurate information
- Listen effectively
- Ask questions
- Acknowledge requests for information
- Openly share ideas
- Attend to non-verbal behaviors

Interpersonal Relations is a broad area that encompasses cooperation and dealing with conflict within the team. Therefore, effective interpersonal relations include working cooperatively with others, working together as opposed to working separately or competitively, and resolving disputes among team members.

Strong interpersonal relations skills are demonstrated by team members who

- Share the work
- Seek mutually agreeable solutions
- Consider different ways of doing things
- Manage/Influence disputes

Group Decision Making/Planning is defined as the ability of a team to gather and integrate information, use logical and sound judgment, identify possible alternatives, select the best solution, and evaluate the consequences.

Strong group decision making and planning skills are demonstrated by team members who work with others to

- Identify problems
- Gather information
- Evaluate information
- Share information
- Understand decisions
- Set goals

Adaptability/Flexibility is defined as the process by which a team is able to use information gathered from the task environment to adjust strategies through the use of compensatory behavior and reallocation of intra-team resources.

Strong adaptability/flexibility skills are demonstrated by team members who

- Provide assistance
- Reallocate tasks
- Provide/Accept feedback
- Monitor/Adjust performance

Core Knowledge Competencies. Regarding the core knowledge competencies, team members must know how and when to use the teamwork skills listed above. Therefore, team knowledge competencies include knowing how to communicate with other team members, how to interact and resolve conflicts, how to plan and make team decisions, and how to adapt and provide assistance to other team members. Such knowledge enables individuals to execute critical teamwork skills and function effectively in a team environment.

The core team knowledge competencies identified above are considered as prerequisites to skill execution. These knowledge competencies are critical components of each team skill (i.e., they comprise the knowledge part of the skill). We present them separately to distinguish what we believe are two critical facets of teamwork: knowing what to do in a team versus doing it. Although the ALL measure focuses on the behavioral alternatives respondents choose in team situations, we believe that knowledge competencies, as defined, are directly related to team member skills and to the level of teamwork achieved.

Core Attitude Competencies. Finally, two attitude competencies were identified: Belief in the Importance of Teamwork and Collective Orientation. These attitudes are brought to the team setting by individuals and can influence the nature of teamwork within a team. As Driskell and Salas (1992) point out, individuals who tend to possess positive attitudes toward teamwork are most likely to take advantage of the benefits teamwork has to offer. Such individuals believe a team approach is better than an individual one; compared to individually-oriented team members, they are better at taking another team member's behavior into account. Each attitude competency is briefly defined below.

Belief in the Importance of Teamwork is defined as the belief that teamwork is critical for successful performance of team tasks.

Collective Orientation is defined as an attraction to, or desire to be part of, a team.

3. A model of teamwork

Based on the literature review and what is generally known about teamwork, Figure 1 presents a model for understanding teamwork for the purposes of ALL. Referring to Figure 1, several things should be noted. First, the skill competencies of Group Decision Making/Planning, Adaptability/Flexibility, and Interpersonal Relations are at the core of teamwork. We believe that team members must know how and when to use these competencies to function effectively within the team. Second, we propose that Communication spans each of the three core areas; it is the glue that holds the team together. For example, Group Decision Making/Planning cannot be accomplished within a team unless team members provide clear and accurate information, listen effectively, and ask questions. Finally, the model proposes that the extent to which an individual is drawn toward teamwork, believes in the importance of teamwork, and has experienced team activity will influence how effectively team skills and behaviors are executed.

Figure 1 also presents a starting point for developing measures for ALL by identifying specific variables to be measured. These include the skills of Group Decision Making/Planning, Adaptability/Flexibility, Interpersonal Relations and Communication, and the attitudes Belief in the Importance of Teamwork and Collective Orientation. Furthermore, Figure 1 presents specific behavioral examples of each skill, as discussed above. These behavioral indicators will be used to construct responses for items measuring teamwork skills. Items that tap respondents' belief in the importance of teamwork and their collective orientation will also be included in the ALL measure.

Figure 1

ALL model for understanding teamwork

Attitudes and experience	Skills		
Attitudes and dispositions	**Group Decision Making/Planning**	**Adaptability/ Flexibility**	**Interpersonal Relations**
	Identify problems	Provide assistance	Share the work
	Gather information	Reallocate tasks	Seek mutually
Experiences	Evaluate information	Provide/Accept	agreeable solutions
	Share information	feedback	Consider different
Implicit theories about teamwork	Understand decisions	Monitor/Adjust	ways of doing things
	Set goals	performance	Manage/Influence disputes
	Communication		
	Provide clear and accurate information		
	Listen effectively		
	Ask questions		
	Acknowledge requests for information		
	Openly share ideasPay attention to non-verbal behaviors		

Prior to discussing our method and approach for developing the Teamwork Scale for ALL, we briefly review the relevant literature on culture. More than other ALL measures, responses to the teamwork measure may be affected by the culture of the respondent. In the next section, we review research that specifically examines the relationship between societal culture and an individual's attitudes, values, beliefs, and behavior in a team. Based on this research and on our understanding of the factors that enhance teamwork, we propose a number of likely relationships that will be demonstrated in the ALL between culture and teamwork.

4. Culture and teamwork

Culture is simply "*the values, beliefs, behavior, and material objects that constitute a people's way of life*" (Macionis, 1993). Research examining the relationship between culture and performance in organizations has tended to focus on people's attitudes, values, beliefs, sources of motivation, and satisfaction and is commonly assumed to predict behavior.

Although alternative categorizations exist (e.g., Trompenaars, 1993), the most commonly used description of cultural comparisons has been developed by Hofstede (1980; 1991). Hofstede conducted the most exhaustive cross-cultural study to date (questionnaire data from 80,000 IBM employees in 66 countries across seven occupations) and established four dimensions of national culture. The four dimensions are the following:

Power Distance: The extent to which the less powerful members of institutions and organizations accept that power is distributed unequally.

Individualism/Collectivism: The extent to which a society is a loosely knit social framework in which people are supposed to take care only of themselves and their immediate families, as opposed to tight social frameworks in which people are integrated into strong cohesive groups that look after them in exchange for loyalty.

Uncertainty Avoidance: The extent to which people feel threatened by ambiguous situations and have created beliefs and institutions that try to avoid them.

Masculinity/Femininity: The extent to which the dominant values in a society tend toward achievement and success and away from caring for others and quality of life.

Research has shown that social dynamics vary according to the norms individuals hold concerning appropriate social behavior and that these norms vary across cultural settings (Triandis, 1989). For example, direct confrontation of one's boss may be acceptable in one culture and avoided in another (Adler, 1986). In fact, preliminary empirical studies have demonstrated large cross-national differences in attitudes regarding task performance across several work domains (Hofstede, 1980; Merritt, 1996; Merritt and Helmreich, 1996). Therefore, it is reasonable to suspect that societal culture exerts important effects on team members' knowledge of acceptable team skills, on members' attitudes toward teamwork, and on team behavior.

4.1 Research on culture and teamwork

Several notable studies have examined the attitudinal differences among workers of different cultures (Evan, 1993). Hofstede (1985) explored a matched sample of employees in a single, multinational corporation in 40 countries. He found wide differences in attitudes toward collaboration. Individualistic countries were more likely to reject collaborative work, preferring to work on their own, whereas collectivist cultures preferred collaborating with others. In related work, Kelly and Reeser (1973) examined the differences between American managers of Japanese ancestry and those of Caucasian ancestry. Similarly, a study by Pizam and Reichel (1977) examined the differences between Israeli managers of Oriental ancestry and those of Western ancestry. In both studies, cultural differences were observed in areas such as respect for formal authority, commitment to long-term employment, paternalism with respect to subordinates, and interest in teamwork.

Cross-national differences in attitudes toward interpersonal interactions have also been found in aviation teams (Helmreich, Merritt, and Sherman, 1996). Current research has demonstrated substantial variability among cultures concerning attitudes toward command responsibility and the captain's role on the flight deck. Cultures differ with respect to members' belief that junior crew members should question the actions of captains. Similarly, individuals from different cultures differ significantly in their endorsement of whether or not they should speak up when they perceive a problem with the flight. Overall, Anglos are more likely than non-Anglos to believe that it is acceptable for crew members to question the captain's decisions, that it is acceptable for the first officer to assume command of the aircraft under certain circumstances, that the captain should not automatically take physical control, and that successful flight deck management depends more than on the captain's individual proficiency.

Parallel findings were found in cross-cultural research, conducted at the Center for Creative Leadership, on teamwork and team leadership. More judicious use of personal prominence and power, greater openness to the ideas and interest of others, and mitigation of tough mindedness are more acceptable among team leaders in Europe, as compared to those in the U.S. (Leslie and Van Velsor, 1998).

Finally, Gibson (1996) found that the relation between team beliefs and team performance differed between American and Indonesian work teams. A collective orientation enhanced team performance, whereas an individualistic orientation inhibited teamwork. Kirkman (1997) found that, in the U.S., Finland, Belgium, and the Philippines, the amount of resistance to working in a team varied, depending upon the cultural orientation of employees. Respondents with individualistic values resisted working in teams more than did respondents with collectivist values. Further, respondents who valued power distance reported higher levels of resistance to self-management than did those who placed a low value on power distance. Currently, Gibson and Zellmer (1997) are engaged in an intercultural analysis on the meaning of teamwork. Although their preliminary results demonstrate that teams have become a pervasive element across the world, the concept of teamwork itself seems to differ as a function of culture.

4.2 Implications for measuring teamwork internationally

Based on the research cited above, it appears that culture can significantly affect the way in which individuals communicate, make decisions, and resolve conflicts in a team. For example, individuals from countries with low power distance (e.g., Austria, Israel, Ireland, and United States) try to minimize inequalities and favor less autocratic leadership and less centralization of authority in teamwork than do individuals from countries with high power distance (e.g., Malaysia, Philippines, Panama, Guatemala, and Puerto Rico). In addition, countries differ significantly in their expression of collectivism, a difference that is likely to affect an individual's desire to participate in teams (i.e., collective orientation) and the extent to which individuals take advantage of the benefits offered by teamwork.

From the standpoint of developing a measure of teamwork for ALL, the research on culture has two important implications. First, although it seems safe to conclude that the core dimensions of teamwork (see Figure 1) generalize to most countries, it also seems likely that the way in which these skills are manifested will vary by nation. For example, communication will be central to teamwork regardless of culture, but team members from different countries may employ somewhat different communication strategies. In an attempt to address this issue, we tried to identify behaviors representing each of the core teamwork skills that were least likely to vary. However, the extent to which we achieved this goal will only be known after testing the teamwork measure in

several different countries. Second, because effective teamwork behaviors likely vary across countries, it may not be possible to construct teamwork items with one "correct" answer. What is considered appropriate team behavior in one country may not be considered appropriate in another. Therefore, our items will attempt to capture information about respondents' knowledge of teamwork across the countries participating in ALL. Norms on these measures will be produced for each country, thereby providing a wealth of information on the nature of teamwork within a country. To the extent that teamwork is manifested differently from culture to culture, cross-cultural comparisons will be neither possible nor appropriate. With these issues in mind, we now turn to a discussion of ALL teamwork measures.

4.3 ALL teamwork measures

The previous sections of this framework have presented our definition of a team and have delineated the core knowledge, skills and attitudes that are associated with effective teamwork (see Figure 1). We have tried to identify individual-level competencies that are generalizable, although we recognize that culture may play a significant role in how individuals express these competencies while functioning in a team.

This section of the framework describes our strategies for assessing teamwork. We first present a series of theoretical and practical assumptions that will guide item development. We present these assumptions here because they have significantly influenced our measurement approach.

4.3.1 Theoretical assumptions

There are four distinguishing features of a team (two or more individuals; a shared or common goal; task interdependence; and a desired productive outcome).

There are generic team competencies held at the individual level that we believe can be measured.

The competencies defined in this framework represent key elements of teamwork that should be measured.

The competencies defined in this framework are critical for successful teamwork.

Attitudes toward teamwork and knowledge of teamwork skills directly affect teamwork.

There are cultural differences associated with teamwork. All cultures will be familiar with the notion of teams, and the competencies reflected in the framework are likely to be common to all cultures. However, these competencies are not necessarily expressed in the same way.

4.3.2 Practical assumptions

Participants will have approximately 30 minutes to complete the Teamwork section of ALL.

Teamwork will be assessed using paper-and-pencil measures.

Although we expect cultural differences in teamwork, we are not trying to measure differences in culture; rather, we emphasize general factors of teamwork with strong cross-cultural relevance.

The same measurement approach will be used to assess teamwork across cultures.

Respondent experience with teams may be work or non-work related (e.g., sports, community, schools, etc.)

Although team processes cannot be directly observed, knowledge about team skills, attitudes toward teamwork, and historical experience with teamwork can be measured.

Among these assumptions, the final practical assumption is most important. As mentioned in the Introduction, it will not be possible to measure respondent team skill competencies directly because the teamwork measure in ALL will be a short paper-and-pencil measure. Measuring team skills has historically required detailed simulations in which team member behaviors are observed and evaluated (D. Baker and Salas, 1992; 1997; Brannick et al., 1997; Ilgen, 1999). Such procedures are inconsistent with the measurement approach of ALL. However, even with these constraints, it is possible to learn a great deal about both the nature of teamwork, and about critical variables that can affect team performance. In particular, respondents' knowledge of teamwork skills (see Figure 1) and respondents' attitudes toward teamwork can be assessed in situation-based items that elicit behavior-oriented, rather than "textbook," responses. Our strategies for measuring each are detailed below.

4.4 Knowledge of teamwork skills

The primary goal of the ALL teamwork measure will be to measure respondent knowledge of teamwork skills, which have been shown to be positively related to team performance (Salas et al., 1999; Stevens and Campion, 1994b). In particular, respondent knowledge of Group Decision Making/Planning, Adaptability/Flexibility, Interpersonal Relations, and Communication will be assessed. Results from this measure will provide information as to how knowledge of teamwork skills is distributed in the adult population within nations.

4.4.1 Measurement approach

In developing our approach for measuring knowledge of teamwork, we faced two significant challenges: (a) because ALL is the first attempt to assess knowledge of teamwork internationally, results from prior research were not available for guidance; and (b) due to practical constraints associated with ALL, the method of measurement was limited to a short paper-and-pencil instrument. Future large-scale assessments of teamwork may consider the use of computer-based simulations or other similar formats to assess team skills more directly (E. Baker, 1998); however, the necessary technology is not currently available to the ALL.

Based on our definition of teamwork, the relevant literature on knowledge tests (Borman, 1991; Dye, Reck and McDaniel, 1993; Hunter, 1986), the domain we sought to measure, and our desire to assess applied knowledge, our questions require respondents to make situational judgments. In personnel selection, both situational judgment questions for written tests and structured interviews have been shown to predict job performance (M. Campion, J. Campion, and Hudson, 1993). Specific to teams, Stevens and Campion (1994b) have reported significant criterion-related validities with supervisory and peer ratings of team performance for a thirty-five-item situational judgment test of teamwork knowledge (although this measure was also significantly correlated with respondent general mental ability). Finally, situational judgment tests have a high degree of face validity for the respondent.

4.4.2 Item development

Initially, an item production grid was constructed to guide item development (refer to Appendix 1.1). The item production grid was derived from the team skill definitions and the behavioral facets representing each skill (i.e., the item production grid in Appendix 1.1 represents the key facets of teamwork in the U.S. and will be modified for different ALL countries). The item production grid is used to ensure that an adequate number of items are developed to cover the skill domains of interest and to specify clearly what each item is intended to measure.

Regarding item construction, short vignettes were initially created. These vignettes describe a fictitious team performing a fictitious team task. Care was taken to ensure that vignettes were based on both work and non-work team situations. Each team described in the vignettes conformed to the definition and characteristics of a "team." To date, five vignettes have been created: one focusing on a toy manufacturing team, one focusing on a marketing team, one focusing on a customer service team and two focusing on community-based teams (one assigned to review school performance and one assigned to clean a park).

Situational judgment items were developed for each vignette. Each item presents a situation, and respondents are asked to rate the effectiveness of each response option on a 5-point scale where 1 indicates "Extremely Bad" and 5 indicates "Extremely Good." To date, eight items have been developed for each vignette, resulting in a total of 40 items. Appendix 1.2 presents several example items. Appendix 1.3 lists all of the items developed thus far.

One issue that was considered, though not specifically accounted for during item development, was the notion of item difficulty. First, unlike other measures included in ALL (i.e., literacy, numeracy, problem solving, etc.), the assessment of teamwork skills (or knowledge of teamwork skills) in the adult population internationally is a new undertaking. Therefore, no research was available to help identify the attributes that might comprise a more difficult and less difficult teamwork item. Certainly, varying the degree to which it is easy to identify the best response from a series of distractors would affect item difficulty. Though this could be done, the ability to respond to more difficult items constructed in this manner would not necessarily reflect more knowledge of teamwork skills. Such responses may be more reflective of a test taker's ability to read, comprehend, and extract the correct information. More importantly, we must acknowledge that the difficulty of teamwork may lie in the execution of team behaviors rather than in the knowledge of what to do. All team members may know what to do in a given team situation, but only the best team members are willing and able to carry out these behaviors in a timely and appropriate fashion that maximizes teamwork. The paper-and-pencil measurement approach used in ALL does not allow for assessing a respondent's skills in terms of actual outcome criteria.

With these issues in mind, we tried to construct items of moderate difficulty. Psychometrically speaking, items of medium difficulty will provide maximum information on the distribution of knowledge of teamwork skills within each ALL country (Crocker and Algina, 1986). Items of medium difficulty were formulated by embedding the "best" alternative for each situational-judgment item ("best" in terms of U.S. research findings) among two alternatives that might reflect other cultures' expressions of team skills and one distracter that virtually no one would be expected to select. In addition, we plan to collect sufficient data during pre-feasibility and feasibility studies to determine each item's difficulty statistically. Item difficulty and other indicators of item performance will be used to select final items for the ALL teamwork measure.

4.4.3 Scoring

Several scoring procedures will be explored during feasibility testing. These range from a Thurstone-like scaling procedure (Anastasi, 1988) in which respondent ratings are compared to country-specific profiles generated for each ALL country to a dichotomous scoring procedure in which each situational judgment item is scored as right or wrong. Whether or not right versus wrong scoring is plausible will be determined by the extent to which rating profiles (i.e., respondent average ratings for the knowledge items) are similar across countries. The final scoring procedure for the team knowledge measure will be selected on the basis of these analyses and practical considerations associated with administration and scoring of ALL.

4.5 Attitudes toward teamwork

Team attitudes are defined as an internal state that influences a team member's choices or decisions to act in a particular way (Cannon-Bowers et al., 1995; Dick and Carey, 1990). Attitudes toward teamwork can have a significant effect on how teamwork skills are actually put into practice. Positive attitudes toward teamwork (Gregorich et al., 1990; Ruffell-Smith, 1979; Helmreich et al.,1986) and an attraction to being part of a team (i.e., collective orientation) have been found to enhance team process and team performance (Driskell and Salas, 1992; Eby and Dobbins, 1997). Therefore, each of these attitudes will be assessed as part of the ALL teamwork measure.

4.5.1 Measurement approach

Unlike the knowledge of teamwork skills, a significant body of work exists on the assessment of attitudes toward teamwork both in the US and internationally (see for example, Eby and Dobbins, 1997; Gregorich et al., 1990; Helmreich et al., 1986). The vast majority of this work, however, has focused on commercial pilot attitudes toward teamwork in the cockpit. Nonetheless, this research provides an excellent starting point for structuring our measurement approach.

A review of past attitude measures indicated that all employed some form of Likert scaling. A similar approach is proposed for ALL. Likert-type scales typically include a series of positive and negative statements about teamwork, and respondents endorse one of a series of graded response options (e.g., strongly agree, agree, neutral, disagree, strongly disagree) for each item. Points are allocated to each response option (e.g., 5 = strongly agree, 4 = agree, etc.) and the sum of these values represent attitude strength.

4.5.2 Item development

Positive and negative statements regarding Belief in the Importance of Teamwork and Collective Orientation were identified and extracted from the research on team attitude measurement (Eby and Dobbins, 1997; Gregorich et al., 1990). Some of these statements were rephrased because they were extracted from a measure designed to assess pilot attitudes toward teamwork in the cockpit. In addition, several new statements were prepared to ensure that a sufficient number of statements were included for reliable measures. In total, 16 statements were developed to measure Belief in the Importance of Teamwork and 15 statements were developed to measure Collective Orientation. Consistent with other approaches, all statements were scaled using a five-point Likert-type scale where 1 = strongly disagree and 5 = strongly agree.

Attitude measures were tested on 192 business students from a mid-western university. Of the cases in which complete demographic data were available, 173 were undergraduate students, 2 were graduate students, and 1 was a professor. The mean age of participants was 22 years old; 74 were female and 108 were male; 73.4% were Caucasian, 12% were Asian or Pacific Islander, 5.2% were African-American and 1% were Hispanic. Most participants (93.8%) also indicated that they had some experience working or participating in a team.

A principal components factor analysis, item-subscale correlations, and a qualitative review of the clarity and potential for cultural bias associated with each item were used to select final items for the two attitude scales. This process resulted in eight items being selected to measure Belief in the Importance of Teamwork (μ = .79) and seven items being selected to measure Collective Orientation (μ = .84). Appendix 1.4 contains the final items for the Team Attitude Scale.

4.5.3 Scoring

The Belief in the Importance of Teamwork Scale and the Collective Orientation Scale will be scored in the same fashion. Total scores will be calculated by summing the points associated with the response alternatives selected by each respondent. Negative statements will be reverse-coded. Scores on the Belief in the Importance of Teamwork Scale can range from a low of 5 to a high of 40, whereas scores on the Collective Orientation Scale can range from a low of 5 to a high of 35. Norms will be developed on the basis of these scale scores, thereby providing information regarding attitudes toward teamwork for countries participating in ALL.

5. Background questionnaire

The Background Questionnaire presents an opportunity to collect information about demographic, social, and economic factors that affect teamwork. Such information should be of interest to policymakers and educators from countries participating in ALL because it will provide information on the determinants of teamwork. Results can be used for structuring policy and/or educational programs to improve the levels of teamwork in the workforce and elsewhere.

Based on what is currently known about teamwork and our approach to measuring teamwork in ALL, we hypothesize that several background variables may have an effect on a respondent's knowledge of team skills, and his or her belief in the importance of teamwork and collective orientation. Specifically, past experience in teams, whether or not the respondent has received formal or informal team training, and demographic variables like respondent age, gender, economic status, and educational level may have an effect. Each of these is briefly discussed in some detail below.

5.1 Experience in teams

The nature and extent of respondents' experiences in teams are likely to significantly affect their attitudes toward teamwork and knowledge of what to do in teams. In addition to including questions about respondents' experiences in teams in the background questionnaire, we developed a short team experience measure (see Appendix 1.5). This scale asks respondents to rate their past experiences in teams on a series of bipolar adjectives. An initial version of this scale was tested on the sample of business students described earlier (refer to item development for the team attitude scales for a detailed description of the sample). Results indicated that the scale was reasonably reliable ($\mu = .79$).

5.2 Team training

Whether or not respondents have received formal or informal team training is likely to have a significant effect on both knowledge of teamwork skills and attitudes toward teamwork. Sufficient research exists to support the efficacy of team training for improving attitudes toward teamwork, increasing knowledge, and enhancing teamwork skills (see for example, Salas et al., 1995; Salas et al., 1999). Collecting information on whether or not ALL respondents have received team training and the nature of training content should prove useful to policymakers interested in improving teamwork skills in the workforce. Data collected through ALL could provide significant insight into specific training strategies that are effective in different countries.

5.3 Demographics

Demographic characteristics such as age, gender, economic status, and educational level may also affect respondent knowledge and attitudes toward teamwork. Little, if any, research on the effects of these variables currently exists. ALL could present an opportunity to assess whether or not knowledge of teamwork skills and attitudes toward teamwork vary as a function of these and other demographic characteristics.

6. Overall summary and conclusions

In closing, this paper presented our framework for assessing teamwork as part of ALL. The framework was derived from the literature on teams and on what is currently known about effective team performance. We have tried to capture the fundamental constructs underlying effective teamwork, in the anticipation that these constructs will generalize to a wide variety of countries, even if their expression may differ across cultures.

In addition to delineating the key facets of teamwork, this paper has also presented our approach to measurement. Although it will not be possible to measure team skill competencies through direct observation, we will nevertheless assess respondents' knowledge of teamwork skills and respondents' attitudes toward working in teams. The results will provide insight into the distribution of these constructs in the international adult population.

Finally, we view this framework as a work in progress for two reasons. First, as with all survey development, we envision conducting significant pilot testing on the approaches we have selected. This testing is likely to lead to revision of our measurement strategies. Although the strategies we have suggested have been effective in other domains, their efficacy for assessing teamwork on an international level has yet to be determined. Second, new information becomes available on teams and the nature of teamwork almost daily. The field is growing and changing concurrently with our efforts. In response, we view our framework as evolving as well; thus, we will incorporate relevant new findings as they become available.

References

Adler, N.J. (1986). *International dimensions of organizational behavior*. Boston, MA: Kent.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan

Ayman, R., and Chemers, M.M. (1983). Relationship of supervisory behavior ratings to work group effectiveness and subordinate satisfaction among Iranian managers. *Journal of Applied Psychology, 68*, 338-341.

Baker, D.P., and Salas, E. (1992). Principles for measuring teamwork skills. *Human Factors, 34,* 469-475.

Baker, D.P., and Salas, E. (1997). Principles for measuring teamwork: A summary and look toward the future. In M.T. Brannick, E. Salas, and C. Prince (Eds.). *Team performance assessment and measurement* (pp. 331-355). Mahwah, NJ: Erlbaum.

Baker, E.L. (1998). *Draft final report for the validation of teamwork skills questionnaire using computer-based teamwork simulations*. University of California, Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Borman, W.C. (1991). Job behavior, performance, and effectiveness. In M.D. Dunnette and L. M. Hough (Eds.) *Handbook of industrial and organizational psychology* (pp. 269-313). Palo Alto, CA: Consulting Psychologists Press.

Boyatzis, R.E. (1982). *The competent manager*. New York: John Wiley and Sons.

Brannick, M.T., Salas, E., and Prince, C. (1997). *Team performance assessment and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brooks, A.K. (1994). Power and the production of knowledge: Collective team learning in work organizations. *Human Resource Development Quarterly, 5*, 213-235.

Campion, M.A., Campion, J.E., and Hudson, J.P. (1994). Structured interviewing: A note on incremental validity and alternate question types. *Journal of Applied Psychology, 79*, 998-1002.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E., and Volpe, C.E. (1995). Defining competencies and establishing team training requirements. In R. Guzzo and E. Salas (Eds.), *Team effectiveness and decision making in organizations*, (pp. 333-380). San Francisco: Jossey Bass.

Cannon-Bowers, J.A., and Salas, E. (1997). A framework for developing team performance measures in training (pp. 45-62). In M.T. Brannick, E. Salas, and C. Prince (Eds.). *Team performance assessment and measurement* (pp. 331-355). Mahwah, NJ: Erlbaum.

Carnevale, A.P., Gainer, I.J., and Meltzer, A.S. (1990). *Workplace basics: The essential skills employers want*. San Francisco, CA: Jossey-Bass.

Cohen, S.G., and Bailey, D.E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management, 23,* 239-290.

Commission on the Skills of the American Workforce. (1990). *America's choice: High skills or low wages!* Rochester, NY: Author.

Conference Board of Canada (1993). *Employability skills profile*. Ottawa, Ont., Canada: Author [The Conference Board of Canada, 255 Smyth Road, Ottawa, Ont. K1H 8M7, Canada, 10/93]

Crocker, L., and Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Reinhart, and Winston.

Dick, W., and Carey, L. (1990). *The systematic design of instruction* (3rd ed.). Glenview, IL: Scott Foresman.

Driskell, J.E., and Salas, E. (1992). Collective behavior and team performance. *Human Factors, 34,* 277-288.

Dye, D.A., Reck, M., and McDaniel, M.A. (1993). Moderators of the validity of written job knowledge measures. *International Journal of Selection and Assessment, 1,* 153-157.

Dyer, J. (1984). Team research and training: A state-of-the-art review. In F.A. Muckler (Ed.), *Human Factors Review: 1984* (pp. 285-323). Santa Monica, CA: Human Factors Society.

Eby, L.T., and Dobbins, G.H. (1997). Collectivistic orientation in teams: An individual and group level analysis. *Journal of Organizational Behavior, 18,* 275-2795.

Gibson, C.G. (1994). The implications of culture for organization structure: An investigation of three perspectives. *Advances in International Comparative Management, 9,* 3-38.

Gibson, C.B. (1996). *They do what they believe they can? Group-efficacy beliefs and group performance across tasks and cultures*. Paper presented at the National Academy of Management meetings, Cincinnati, OH.

Gibson, C.B., and Zellmer, M.E. (1997). *Intercultural analysis of the meaning of teamwork: Evidence from six multinational corporations*. Paper presented at the Academy of International Business Conference, Monterrey, Mexico.

Gordon, J. (1992). Work teams: How far have they come? *Training*, (October 1992), 59-65.

Gregorich, S.E., Helmreich, R.L., and Wilhelm, J.A. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology, 75,* 682-690.

Guzzo, R.A., and Shea, G.P. (1992). Group performance and inter-group relations in organizations. In M.D. Dunnette and L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., 3, 269-313). Palo Alto, CA: Consulting Psychologists Press, 1992.

Hackman, J.R. (1987). The design of work teams. In J.W. Lorsch (Ed.). *Handbook of Organizational Behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice Hall.

Helmreich, R.L., Foushee, H.C., Benson, R., and Russini, W. (1986). Cockpit management attitudes: Exploring the attitude-behavior linkage. *Aviation, Space, and Environmental Medicine, 57,* 1198-1200.

Helmreich, R.L., Merritt, A.C., and Sherman, P.J. (1996). Human factors and national culture. *International Civil Aviation Organization Journal, 51*, 14-16.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values.* Beverly Hills: Sage.

Hofstede, G. (1985). The interaction between national and organizational value systems, *Journal of Management Studies, 22,* 347-357.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind.* UK: McGraw-Hill.

Hunter, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29,* 340-362.

Ilgen, D.R. (1999). Teams embedded in organizations. *American Psychologist, 54,* 129-139.

Kirkman, B.L. (1997). The impact of cul-tural values on employee resistance to teams: Toward a model of globalized self-managing work team effectiveness. A*cad-emy of Management Review, 22*, 730-757.

Leslie. J., and Van Velsor, E. (1998). *A Cross-National Comparison of Effective Leadership and Teamwork.* Greensboro, NC: Center for Creative Leadership.

Macionis, J.J. (1997*). Sociology* (6th ed.) Upper Saddle River, NJ: Prentice Hall.

McGrath, J.E. (1984). *Groups: Interaction and performance.* Englewood Cliffs, NJ: Prentice Hall.

McGrath, J. (1997). Small group research, that once and future field: An interpretation of the past with an eye to the future. *Group Dynamics: Theory, Research, and Practice, 1(1)*, 7-27.

McIntyre, R.M., and Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. Guzzo, E. Salas, and Associates (Eds.), *Team effectiveness and decision making in organizations,* (pp. 9-45). San Francisco, CA: Jossey-Bass.

Merritt, A.C. (1996). *National culture and work attitudes in commercial aviation: A cross-cultural investigation.* Unpublished doctoral dissertation. The University of Texas at Austin.

Merritt A.C., and Helmreich, R.L. (1996-a). Human factors on the flight deck: The influence of national culture. *Journal of Cross-Cultural Psychology, 27*, 5-24.

Morgan, B.B., Glickman, A.S., Woodward, E.A., Blaiwes, A., and Salas, E. (1986). *Measurement of team behaviors in a Navy environment* (NTSC Report No. 86-014). Orlando, FL: Naval Training Systems Center.

O'Neil, H.F., Chung, G.K.W.K., Brown, R.S. (1997). Use of network simulations as a context to measure team competencies. In H.F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Oser, R.L., McCallum, G.A., Salas, E., and Morgan, B.B., Jr. (1992). Toward a definition of teamwork: Behavioral elements of successful teams (*NTSC Technical Report No. 89-018*). Orlando, FL: Naval Training Systems Center.

Parry, S.B. (1998). Just what is a competency (And why should we care?). *Training*, 58-64.

Pizam, A., and Reichel, A. (1977). Cultural determinants of managerial behavior. *Management International Review, 17,* 65-72.

Ruffell-Smith, H.P. (1979). A simulator study of the interaction of pilot workload with errors (NASA Technical Report No. TM-78482). Moffett Field, CA: National Aeronautics and Space Administration-Ames Research Center.

Salas, E., Bowers, C.A., and Cannon-Bowers, J.A. (1995). Military team research: 10 years of progress. *Military Psychology, 7,* 55-75.

Salas, E., Dickinson, T., Converse, S.A., and Tannenbaum, S.I. (1992). Toward an understanding of team performance and training. In R.W. Swezey and E. Salas (Eds.), *Teams: Their training and performance*, (pp. 3-29). Norwood, NJ: Ablex.

Salas, E., Fowlkes, J.E., Stout, R.J., Milonvich, D.M., and Prince, C. (1999). *Does CRM training improve teamwork skills in the cockpit? Two evaluation studies.* Submitted for publication.

Stevens, M.J., and Campion, M.A. (1994a). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management, 20,* 503-530.

Stevens, M.J., and Campion, M.A. (1994b). Staffing teams: Development and validation of the teamwork KSA test. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Nashville, TN.

Triandis, H.C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review, 96,* 506-520.

Trompenaars, F. (1994). *Riding the waves of culture: Understanding diversity in global business*. Chicago, IL: Irwin Professional Publishing.

US Department of Labor (1991, June). *What work requires of schools: A SCANS report for American 2000*. Washington, DC: US Department of Labor, The Secretary's Commission of Achieving Necessary Skills (SCANS).

US Department of Labor (1992a). *Learning and living: A blueprint for high performance*. Washington, DC: US Department of Labor, The Secretary's Commission of Achieving Necessary Skills (SCANS).

US Department of Labor (1992b). *Skills and tasks for jobs: A SCANS report for American 2000.* Washington, DC: US Department of Labor, The Secretary's Commission of Achieving Necessary Skills (SCANS).

Vaziri, M.T., Lee, J.W., and Krieger, J.L. (1988). Onda Moku: The true pioneer of management through respect for humanity. *Leadership and Organization Development Journal, 9,* 3-7.

Appendices

Appendix 1.1

Item Production Grid

Teamwork Skill	Behavioral Requirements	Items
Group Decision Making/Planning	Identify problems	
	Gather information	
	Evaluate information	
	Share information	
	Understand decisions	
	Set goals	
Adaptability/Flexibility	Provide assistance	
	Reallocate tasks	
	Provide/Accept feedback	
	Monitor/Adjust performance	
Interpersonal Relations	Share the work	
	Seek mutually agreeable solutions	
	Consider different ways of doing things	
	Manage/Influence disputes	
Communication	Provide clear and accurate information	
	Listen effectively	
	Ask questions	
	Acknowledge requests for information	
	Openly share ideas	
	Pay attention to non-verbal behavior	

Appendix 1.2

Example Knowledge Items

The following survey describes a team and several situations that the team encounters. After each situation, there are several response options describing what the team could do. For each option listed, rate the quality of the option on the following 1-to-5 scale.

Rating Scale

1	2	3	4	5
Extremely Bad	Somewhat Bad	Neither Bad Nor Good	Somewhat Good	Extremely Good

Vignette 4

A team of volunteers cleans a community park each month. The park is so large that the team needs an entire day to clean it.

Item 1

Members of the team have always worked well together. Recently, the community requested that the park be cleaned more often. The team meets to discuss this requirement, but team members disagree about how to proceed. To help this situation, team members should:

a) _____ Act as though the differences in opinion are not very important.

b) _____ Write down the various opinions about how to proceed and have a team member select one at random.

c) _____ Ask someone from outside the team to act as a mediator at the next meeting.

d) _____ Conduct a candid discussion about the issues on which the team members disagree.

Item 2

The team is asked to periodically rake all the leaves in the park every few weeks during the fall. This situation places a new demand on the team. To cope with this increased demand on its time, the team should:

a) _____ Refuse to do the additional work.

b) _____ Distribute the additional work equally among team members.

c) _____ Assign the additional work to the newest team member.

d) _____ Ask another team to do half the work.

Item 3

One team member leaves the team and a new individual volunteers. The next month the park is cleaned, the team should:

a) _____ Assign the new team member his fair share of the work, but be willing to help out, if necessary.

b) _____ Assign the new team member only the easiest tasks.

c) _____ Encourage the new team member to learn the work by trial and error.

d) _____ Tell the new team member to stay out of the way and watch what the other team members are doing.

Item 4

No one on the team wants to clean the park bathrooms. To resolve this situation, the team should:

a) _____ Decide through a lottery who cleans the bathrooms each time.

b) _____ Have the newest team member clean the bathrooms.

c) _____ Rotate the responsibility of cleaning the bathrooms to a different team member each month.

d) _____ Refuse to clean the bathrooms, since no one on the team wants to do it.

Item 5

The team is requested to make a recommendation on how to improve the park. When the team meets to decide on its recommendation, the team should:

a) _____ Discuss a wide variety of recommendations before making a decision.

b) _____ Allow each team member to suggest one recommendation for consideration by the team.

c) _____ Assign the responsibility for making a recommendation to the team member who seems to know the most about parks.

d) _____ Tell the community it is not the team's job to make a recommendation.

Item 6

The next park cleaning is scheduled for a holiday and most team members will be out of town. The team meets to reschedule cleaning the park. During this meeting, team members should:

a) _____ Try to participate as much as possible in the decision making process.

b) _____ Hide their own feelings to promote good relationships.

c) _____ Anticipate and discuss potential problems with cleaning the park on a different day.

d) _____ Encourage quieter team members to go along with the most outspoken members in order to reach a quick decision.

Item 7

While cleaning the park, a team member is uncertain about what another team member has asked him to do. The team member should:

a) _____ Try to guess what the other team member wanted.

b) _____ Ignore the request; the other team member will ask again if it's important.

c) _____ Ask the other team member to repeat what he or she said.

d) _____ Tell the other team member to speak more clearly.

Appendix 1.3

Teamwork Situational Judgment Items

Teamwork Skill	Behavioral Requirements	Items
Group Decision Making/ Planning	Identify problems	V4-I6
	Gather information	V3-I1
	Evaluate information	V1-I2; V1-I8; V2-I6; V3-I2
	Share information	V4-I5; V5-I1
	Understand decisions	V3-I6; V5-I5
	Set goals	V2-I2; V5-I2
Adaptability/Flexibility	Provide assistance	V1-I1; V4-I3
	Reallocate tasks	V2-I4; V4-I2
	Provide/Accept feedback	V1-I7; V3-I4; V5-I4
	Monitor/Adjust performance	V2-I5; V3-I3; V5-I3
Interpersonal Relations	Share the work	V4-I4
	Seek mutually agreeable solutions	V1-I3; V1-I4
	Consider different ways of doing things	V2-I1; V5-I6
	Manage/Influence disputes	V2-I3; V3-I5; V4-I1
Communication	Provide clear and accurate information	V3-I8
	Listen effectively	V2-I6; V4-I8
	Ask questions	V4-I7; V5-I7
	Acknowledge requests for information	V1-I5
	Openly share ideas	V2-I7; V2-I8
	Pay attention to non-verbal behavior	V3-I7; V5-I8

Note: V2-I8 indicates Vignette 2 — Item 8.

Appendix 1.4

Team Attitude Scale

For each item, please indicate your response by circling the appropriate number for each item in the scale below.

	Strongly disagree		Neither agree or disagree		Strongly agree
1. Teamwork skills deserve more attention in the workplace.	1	2	3	4	5
2. Teams make better decisions than individuals.	1	2	3	4	5
3. Given a choice, I would rather work alone than do a job where I have to work in a team.	1	2	3	4	5
4. It is impossible to function in today's society without being a good team player.	1	2	3	4	5
5. I prefer to participate in team-oriented activities.	1	2	3	4	5
6. Teams always outperform individuals.	1	2	3	4	5
7. Everyone should be taught to be a good team player.	1	2	3	4	5
6. Teams always outperform individuals.	1	2	3	4	5
7. Everyone should be taught to be a good team player.	1	2	3	4	5
8. I prefer to work on teams where team members perform their own tasks independently rather than working together.	1	2	3	4	5
9. I find that working as a member of a team increases my ability to perform effectively.	1	2	3	4	5
10. I find working in a team to be very satisfying.	1	2	3	4	5
11. Teamwork is one of the most important skills in life.	1	2	3	4	5
12. I prefer to be rewarded for my team's performance rather than my individual performance.	1	2	3	4	5
13. People with strong teamwork skills will always be successful.	1	2	3	4	5
14. Teams plan better than individuals.	1	2	3	4	5
15. I prefer working as part of a team to working alone.	1	2	3	4	5

Appendix 1.5

Team Experience

For each word pair, please assess your overall past experience across all of the teams you have participated in by circling the appropriate number on the scale provided.

Note: If you have never worked/participated in a team, please provide your perceptions as to what you think working/participating in most teams would be like.

Example Items

		Neutral
1. Competitive	1 ——— 2 ——— 3 ——— 4 ——— 5 ——— 6 ——— 7	Cooperative
		Neutral
2. Open	1 ——— 2 ——— 3 ——— 4 ——— 5 ——— 6 ——— 7	Closed
		Neutral
3. Rigid	1 ——— 2 ——— 3 ——— 4 ——— 5 ——— 6 ——— 7	Flexible
		Neutral
4. Trusting	1 ——— 2 ——— 3 ——— 4 ——— 5 ——— 6 ——— 7	Distrustful
		Neutral
5. United	1 ——— 2 ——— 3 ——— 4 ——— 5 ——— 6 ——— 7	Divided

Appendix 2.1

Frequencies for Canada and Italy for the teamwork modules during the feasibility studies

Section L Teamwork - Past Experience – Frequencies – Canada only

Below you will find a list of various types of teams along with a description of each type and some examples.

Please read the descriptions and examples and then indicate the context in which you have ever participated in any of the types of teams. Then, tell us how much experience you have had in each of the teams in which you participated.

Type of Team and Description	Did you participate in this type of team in a work organization?	Did you participate in this type of team in a student organization?	Did you participate in this type of team in a volunteer or community organization?	Did you participate in this type of team in a religious organization?
L11	**L1BA L1AA**	**L1D L1CA**	**L1F L1EA**	**L1HA L1G**
Management team Team is responsible for coordinating, budgeting, and staffing several units within an organization. **Examples:** Executive and Management teams	Yes [1] 30 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 70	Yes [1] 15 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 84	Yes [1] 29 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 70	Yes [1] 8 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 91
	→ **Go to question L1C**	→ **Go to question L1E**	→ **Go to question L1G**	→ **Go to question L2 next type of team**
L2	**L2BA L2AA**	**L2DA L2CA**	**L2F L2EA**	**L2HA L2G**
Project Teams (task forces) Team is responsible for accomplishing a specific set of tasks (develop and carry out an idea, plan, etc.) within a given period of time. The team then disbands when the task is complete. **Examples:** New-product teams, design teams	Yes [1] 36 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 64	Yes [1] 25 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 74	Yes [1] 31 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 69	Yes [1] 7 → How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 92
	→ **Go to question L2C**	→ **Go to question L2E**	→ **Go to question L2G**	→ **Go to question L3 next type of team**

Appendix 2.1 - continued

Frequencies for Canada and Italy for the teamwork modules during the feasibility studies

Below you will find a list of various types of teams along with a description of each type and some examples.

Please read the descriptions and examples and then indicate the context in which you have ever participated in any of the types of teams. Then, tell us how much experience you have had in each of the teams in which you participated.

Type of Team and Description	Did you participate in this type of team in a work organization?	Did you participate in this type of team in a student organization?	Did you participate in this type of team in a volunteer or community organization?	Did you participate in this type of team in a religious organization?
L31	**L3BA L3AA**	**L3DBA L3CA**	**L3FA L3EAA**	**L3HBA L3G**
Production teams Team is responsible for repeatedly producing a specific product. **Examples:** assembly teams, coal-mining crews	Yes [1] 22 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 5 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 9 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 5 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years
	No [2] 79 → **Go to question L3C**	No [2] 94 → **Go to question L3E**	No [2] 90 → **Go to question L3G**	No [2] 94 → **Go to question L4 next type of team**
L41	**L4BA L4AA**	**L4DA L4CA**	**L4FA L4EA**	**L4H L4G**
Service Teams Team is responsible for repeatedly providing a specific service. Repeated transactions with customers or clients. **Examples:** Retail sales teams, maintenance teams, airline attendent crews, customer service teams	Yes [1] 53 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 7 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 17 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years	Yes [1] 3 → How much experience have you had in this type of team? [1] less than 1 year [2] 1 year [3] 2 - 3 years [4] 4 - 5 years [5] more than 5 years
	No [2] 48 → **Go to question L4C**	No [2] 93 → **Go to question L4E**	No [2] 84 → **Go to question L4G**	No [2] 97 → **Go to question L5 next type of team**

Appendix 2.1 - continued

Frequencies for Canada and Italy for the teamwork modules during the feasibility studies

Below you will find a list of various types of teams along with a description of each type and some examples.

Please read the descriptions and examples and then indicate the context in which you have ever participated in any of the types of teams. Then, tell us how much experience you have had in each of the teams in which you participated.

Type of Team and Description	Did you participate in this type of team in a work organization?	Did you participate in this type of team in a student organization?	Did you participate in this type of team in a volunteer or community organization?	Did you participate in this type of team in a religious organization?
L5	**L5BA L5AA**	**L5DA L5CA**	**L5FA L5EA**	**L5H L5GA**
Action and Performing Teams Team is responsible for performing a specific event within a given period of time. **Examples:** Patient care teams, cockpit crews, firefighting teams, rescue teams, bands or musician ensembles, performing arts ensembles	Yes [1] 21 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 81 ➜ **Go to question L5C**	Yes [1] 12 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 88 ➜ **Go to question L5E**	Yes [1] 21 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 80 ➜ **Go to question L5G**	Yes [1] 3 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 95 ➜ **Go to question L6 next type of team**
L6	**L6BA L6AA**	**L6DA L6CA**	**L6FA L6EA**	**L6H L6GA**
Parallel Teams Brought together to generate ideas, make suggestions, recommendations, or to solve a specific problem. **Examples:** Ad hoc committees. Quality Circles, TQM	Yes [1] 29 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 69 ➜ **Go to question L6C**	Yes [1] 16 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 83 ➜ **Go to question L6E**	Yes [1] 14 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 84 ➜ **Go to question L6G**	Yes [1] 3 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 95 ➜ **Go to question L7 next type of team**

Appendix 2.1 -cconcluded

Frequencies for Canada and Italy for the teamwork modules during the feasibility studies

Below you will find a list of various types of teams along with a description of each type and some examples.

Please read the descriptions and examples and then indicate the context in which you have ever participated in any of the types of teams. Then, tell us how much experience you have had in each of the teams in which you participated.

Type of Team and Description	Did you participate in this type of team in a work organization?	Did you participate in this type of team in a student organization?	Did you participate in this type of team in a volunteer or community organization?	Did you participate in this type of team in a religious organization?
L77	L7BA L7AA	L7DA L7CA	L7FA L7EA	L7H L7GA
Other Teams **Examples:** Boy or Girl Scouts, 4-H clubs, future Farmers of America, academic clubs (e.g., school newspaper, yearbook, science club), school spirit or fund raising clubs, student council/student government	Yes [1] 21 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 79	Yes [1] 30 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 70	Yes [1] 36 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 63	Yes [1] 5 ➜ How much experience have you had in this type of team? [1] ◯ less than 1 year [2] ◯ 1 year [3] ◯ 2 - 3 years [4] ◯ 4 - 5 years [5] ◯ more than 5 years No [2] 92

In the last 12 months, were you involved in any team sports (e.g., soccer, basketball, etc.)?

L8A

1 ○ Yes ➜ **Go to question L8A**
2 ○ No ➜ **Go to question L9**

Did you ... (MARK ALL THAT APPLY).

L8A

		Yes	No
1 ○	Play in an organized league with other teams	16	8
2 ○	Play on an informal basis with friends and colleagues	3	8
3 ○	Serve as a coach in an organized youth/adult league	0	0
4 ○	Serveas a referee in an organized youth/adult league	0	0

How frequently did you play, coach or referee?

L8B

1 12	Two or more times per week
2 7	Once per week
3 0	Once per month
4 0	Two to three times per year

In the last 12 months, how would you describe the extent of your involvement in volunteer, community and/or religious organizations? (MARK ALL THAT APPLY).

L9

1 7	active leader (for example, serve as an officer)
2 19	active member (for example, attends all club meetings and always helps out when needed)
3 23	member - I do some work on occasion
4 5	member - but I=m not really active
5 47	Do not participate in such organizations

In your current job, approximately what percentage of your time at work do you spend working in a team? ***(If you have more than one job, tell us about the one at which you work the most hours.)***

L10

1 20	I do not work on a team
2 9	Less than 20%
3 5	21 B 40%
4 4	41 B 60%
5 8	61 B 80%
6 17	More than 80%
N 39	I do not have a job

In your current job, how important is teamwork for accomplishing your tasks?

L11

[1] 12 I do not work on a team
[2] 2 Not Important
[3] 12 Somewhat Important
[4] 7 Important
[5] 18 Very Important
[6] 13 Extremely important
[N] 37 I do not have a job

For the following statements below, mark all that apply.

L12

[1] ◯ In the last 12 months, I have taken a formal course(s) on teamwork from an educational institution (school, college, university, institute, etc.)
[2] ◯ In the last 12 months, I participated in a training program(s) on teamwork provided by my employer
[3] ◯ In the last 12 months, I have received on-the-job training on teamwork from my employer.
[4] ◯ In the last 12 months, I have read a book on teamwork.

In the last 12 months, please indicate the types of teams (either at work or outside of work) you have been a member of (mark all that apply).

L130

[1] ◯ I have not been a member of a team during the last 12 months.
[2] ◯ I have worked on a team where team members reported to or took direction from a higher-ranking individual on the team.
[3] ◯ I have worked on a team where all members were basically the same rank and one person was elected or appointed as the leader.
[4] ◯ I have worked on a team where all members were basically the same rank and we shared responsibilities for coordinating activities.
[5] ◯ I have worked on a team where I was dependent upon other team members doing their job in order for me to do mine.
[6] ◯ I have worked on a team where I could only meet my goals if other team members met theirs.

M1

Organizations (both work and non-work) throughout the world are increasingly relying on work teams. A work team is defined as a group of individuals who produce goods or services for which they are all accountable. A distinguishing feature of all work teams is that members of the team are "interdependent", that is, members of the team must work together to perform work tasks. No team member can accomplish the team's tasks alone. Below is a series of statements about work teams that are associated with specific things that teams do, such as making decisions, communicating, interacting, etc. Please indicate the extent to which you agree or disagree with each statement.

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1)	Teams should always choose the first suggestion offered when trying to make an important decision.	[1] 92	[2] 38	[3] 16	[4] 6	[5] 3
2)	Teams should gather information from a wide variety of sources when making an important decision.	[1] 0	[2] 0	[3] 2	[4] 48	[5] 102
3)	Team members should try to anticipate potential problems with the team's decision.	[1] 0	[2] 1	[3] 15	[4] 76	[5] 59
4)	Teams should let the most skilled team member make important decisions for the rest of the team.	[1] 34	[2] 60	[3] 30	[4] 18	[5] 8
5)	Team members should ignore most team decisions.	[1] 100	[2] 42	[3] 3	[4] 3	[5] 1
6)	Team members should discuss potential problems with the team's decision.	[1] 1	[2] 1	[3] 9	[4] 68	[5] 70
7)	Teams should assign one person from the team to make the decision for the rest of the team.	[1] 76	[2] 51	[3] 12	[4] 9	[5] 1
8)	Teams should make important decisions after evaluating different alternatives.	[1] 1	[2] 1	[3] 6	[4] 73	[5] 68
9)	Team members should quit the team if they don't like a decision.	[1] 43	[2] 56	[3] 36	[4] 9	[5] 3
10)	Team members should refrain from voicing their opinions about team decisions.	[1] 62	[2] 52	[3] 17	[4] 14	[5] 3
11)	Team members should try to understand the reasons for the team's decision.	[1] 1	[2] 12	[3] 14	[4] 84	[5] 49
12)	Teams should ask someone who is not a member of the team to make an important decision.	[1] 63	[2] 54	[3] 25	[4] 5	[5] 2
13)	Teams should gather information mainly from the team's supervisor when making a decision.	[1] 32	[2] 57	[3] 33	[4] 20	[5] 5
14)	Team members should share information when making an important decision.	[1] 1	[2] 1	[3] 2	[4] 55	[5] 90
15)	Teams members should ignore the schedule and perform the work at their own pace.	[1] 42	[2] 66	[3] 27	[4] 10	[5] 3
16)	Team members should help other members with the work if they need it.	[1] 0	[2] 2	[3] 6	[4] 84	[5] 57
17)	Teams should try to get everyone on the team to work at the same pace.	[1] 6	[2] 37	[3] 37	[4] 53	[5] 15
18)	Teams should revise deadlines for team members who are behind schedule.	[1] 6	[2] 46	[3] 42	[4] 46	[5] 8
19)	Teams should tell new members to stay out of the way.	[1] 75	[2] 63	[3] 6	[4] 2	[5] 2
20)	Teams should ask individuals outside of the team how the team is doing.	[1] 12	[2] 25	[3] 44	[4] 57	[5] 9
21)	Teams should make the person who shows up last do any additional work.	[1] 58	[2] 65	[3] 19	[4] 4	[5] 2
22)	Team members should make suggestions to other members as to how to improve their performance.	[1] 4	[2] 4	[3] 23	[4] 81	[5] 36
23)	Teams should assign new team members the easy tasks.	[1] 18	[2] 53	[3] 40	[4] 34	[5] 3
24)	Teams should distribute new work equally among team members.	[1] 1	[2] 4	[3] 17	[4] 82	[5] 44
25)	Team members should evaluate each other's performance.	[1] 5	[2] 26	[3] 38	[4] 61	[5] 17
26)	Teams should punish members who make mistakes.	[1] 60	[2] 60	[3] 18	[4] 8	[5] 1
27)	Team members should act as though differences of opinion are not very important.	[1] 30	[2] 47	[3] 17	[4] 43	[5] 12
28)	Team members should be open to different ways of doing things.	[1] 1	[2] 0	[3] 3	[4] 76	[5] 68
29)	Teams members should be open to suggestions.	[1] 0	[2] 2	[3] 2	[4] 82	[5] 63
30)	Teams should discourage team members from bringing up differences of opinion.	[1] 44	[2] 77	[3] 14	[4] 8	[5] 3
31)	Team members should distract other members during team meetings.	[1] 87	[2] 49	[3] 6	[4] 1	[5] 2

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
32)	Teams should ask someone outside the team to act as a mediator at team meetings.	[1] 23	[2] 45	[3] 53	[4] 24	[5] 2
33)	Team members should discuss other team members' behaviour without them present.	[1] 64	[2] 55	[3] 18	[4] 8	[5] 1
34)	Team members should treat other team members with dignity and respect.	[1] 0	[2] 1	[3] 1	[4] 40	[5] 105
35)	Team members should conduct candid discussions about issues when they disagree.	[1] 3	[2] 6	[3] 16	[4] 67	[5] 56
36)	Team members should express anger toward members who disagree with the rest of the team.	[1] 78	[2] 56	[3] 8	[4] 3	[5] 1
37)	Team members should try to find common ground when they disagree.	[1] 3	[2] 5	[3] 9	[4] 82	[5] 48
38)	Team members should make jokes about other members' behaviour.	[1] 58	[2] 51	[3] 27	[4] 9	[5] 1
39)	Teams should discipline team members who disagree with the rest of the team.	[1] 56	[2] 51	[3] 12	[4] 21	[5] 4
40)	Team members should only let the most experienced team members talk.	[1] 71	[2] 58	[3] 12	[4] 5	[5] 0
41)	Team members should pay attention to other team members' tone of voice.	[1] 3	[2] 24	[3] 40	[4] 66	[5] 12
42)	Team members should ask questions of other team members.	[1] 4	[2] 7	[3] 15	[4] 78	[5] 41
43)	Teams members should try to guess what other team members are going to say.	[1] 40	[2] 56	[3] 42	[4] 6	[5] 2
44)	Team members should ignore other member's requests.	[1] 73	[2] 60	[3] 4	[4] 6	[5] 2
45)	Team members should always speak in a specific order.	[1] 23	[2] 57	[3] 33	[4] 23	[5] 8
46)	Team members should only share information that team members must know to do their jobs.	[1] 24	[2] 66	[3] 23	[4] 26	[5] 8
47)	Team members should provide clear and accurate information to one another.	[1] 0	[2] 1	[3] 6	[4] 76	[5] 63
48)	Team members should openly share ideas, opinions, and problems.	[1] 3	[2] 1	[3] 6	[4] 76	[5] 63
49)	Team members should ignore each other's suggestions.	[1] 80	[2] 55	[3] 7	[4] 3	[5] 1
50)	Team members should respond to other team members' questions.	[1] 3	[2] 3	[3] 8	[4] 90	[5] 42
51)	Team members should not ask other team members to repeat what they said.	[1] 38	[2] 69	[3] 26	[4] 10	[5] 1
52)	Team members should spend time talking about activities outside of work.	[1] 15	[2] 19	[3] 60	[4] 41	[5] 10

→ **PLEASE GO TO SECTION N**

N1

Below is a series of statements about team attitudes. Please indicate the extent to which you agree or disagree with each statement.

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1)	Teamwork deserves more attention in the workplace.	[1] 2	[2] 8	[3] 22	[4] 79	[5] 37
2)	Teams make better decisions than individuals.	[1] 5	[2] 13	[3] 42	[4] 64	[5] 25
3)	Given a choice, I would rather work alone than do a job where I have to work in a team.	[1] 17	[2] 52	[3] 41	[4] 23	[5] 14
4)	It is impossible to function in today's world without being a good team player.	[1] 5	[2] 16	[3] 39	[4] 66	[5] 24
5)	I prefer to participate in team-oriented activities.	[1] 2	[2] 23	[3] 49	[4] 56	[5] 18
6)	Teams always outperform individuals.	[1] 16	[2] 42	[3] 60	[4] 18	[5] 12
7)	Everyone should be taught to be a good team player.	[1] 0	[2] 8	[3] 27	[4] 79	[5] 35
8)	I prefer to work on teams where team members perform their own tasks independently rather than working together.	[1] 12	[2] 33	[3] 54	[4] 37	[5] 11
9)	I find that working as a member of a team increases my ability to perform effectively.	[1] 2	[2] 14	[3] 23	[4] 87	[5] 20
10)	I find working in a productive team to be very satisfying.	[1] 0	[2] 4	[3] 23	[4] 83	[5] 37
11)	Teamwork is one of the most important skills in life.	[1] 1	[2] 13	[3] 30	[4] 73	[5] 31
12)	Teams should ask someone who is not a member of the team to make an important decision.	[1] 63	[2] 54	[3] 25	[4] 5	[5] 2
13)	Teams should gather information mainly from the team's supervisor when making a decision.	[1] 32	[2] 57	[3] 33	[4] 20	[5] 5
14)	Team members should share information when making an important decision.	[1] 1	[2] 1	[3] 2	[4] 55	[5] 90
15)	Teams members should ignore the schedule and perform the work at their own pace.	[1] 42	[2] 66	[3] 27	[4] 10	[5] 3

→ **PLEASE GO TO SECTION O**

Appendix 2.2

Results of feasibility study on Team Attitudes Measure

Table 1

Correlations, descriptive statistics, and reliability data for team attitudes scale (combined sample)

Subscale	Belief in teams	Collective orientation	Mean	Standard deviation	N
Belief in teams	1.00		28.45	4.94	141
Collective orientation	0.67	1.00	23.96	4.57	142
Alpha	0.82	0.81			

Note: Alpha for the 15-item measure was 0.88.

Table 2

Correlations, descriptive statistics, and reliability data for team attitudes scale (Canadian sample)

Subscale	Belief in teams	Collective orientation	Mean	Standard deviation	N
Belief in teams	1.00		28.64	4.92	99
Collective orientation	0.61	1.00	24.44	4.65	99
Alpha	0.81	0.82			

Note: Alpha for the 15-item measure was 0.88.

Table 3

Correlations, descriptive statistics, and reliability data for team attitudes scale (Italian sample)

Subscale	Belief in teams	Collective orientation	Mean	Standard deviation	N
Belief in teams	1.00		28.00	5.01	42
Collective orientation	0.67	1.00	22.84	4.20	43
Alpha	0.84	0.76			

Note: Alpha for the 15-item measure was 0.89.

Team Non-Cognitive Measure

Results

Note: The results below are based on the 24 "best" items that we identified from the pool of 52 items that were administered.

Table 4

Correlations, descriptive statistics, and reliability data for non-cognitive measure (Combined sample)

Scale	DM	AF	IR	CM	Mean	Standard deviation	N
Decision making	1.00				26.15	2.79	144
Adaptability/Flexibility	0.48	1.00			24.49	2.92	147
Interpersonal relations	0.54	0.52	1.00		26.20	2.91	143
Communication	0.56	0.52	0.67	1.00	26.08	3.08	142

Note: Alpha for the 24-item measure was 0.88.

Table 5

Correlations, descriptive statistics, and reliability data for non-cognitive measure (Canadian sample)

Scale	DM	AF	IR	CM	Mean	Standard deviation	N
Decision making	1.00				26.61	2.79	101
Adaptability/Flexibility	0.48	1.00			24.80	2.99	101
Interpersonal relations	0.51	0.50	1.00		26.35	3.01	98
Communication	0.52	0.51	0.66	1.00	26.70	2.89	99

Table 6

Correlations, descriptive statistics, and reliability data for non-cognitive measure (Italian sample)

Scale	DM	AF	IR	CM	Mean	Standard deviation	N
Decision making	1.00				25.07	2.52	144
Adaptability/Flexibility	0.40	1.00			23.80	2.65	147
Interpersonal relations	0.63	0.54	1.00		25.89	2.68	143
Comm.	0.52	0.47	0.74	1.00	24.63	3.02	142
Alpha	0.66	0.50	0.68	0.66			

Note: Alpha for the 24-item measure was 0.86.

The following results are reported for the non-cognitive measure and attitude measure as a function of respondent experience in different types of work teams. Data are based on the Canadian sample only. Means denoted by (*) are significantly different at $p<.05$. No significant effects were found for the variables: management teams, parallel teams, action and performing teams, and other teams.

Table 7

Project teams

Scale	Experience	N	Means
Decision making	Yes	36	27.25
	No	63	26.28
Adaptability	Yes	36	25.00
	No	63	24.68
Interpersonal	Yes	34	26.62
	No	62	26.19
Communication	Yes	35	27.63 *
	No	62	26.19 *
Belief in teams	Yes	34	28.76
	No	63	28.60

Table 8

Production teams

Scale	Experience	N	Means
Decision making	Yes	22	27.73 *
	No	78	26.28 *
Adaptability	Yes	22	24.31
	No	78	24.91
Interpersonal	Yes	21	27.19
	No	76	26.12
Communication	Yes	20	27.90 *
	No	78	26.41 *
Belief in teams	Yes	21	28.90
	No	77	28.66

Table 9

Service teams

Scale	Experience	N	Means
Decision making	Yes	53	27.08
	No	47	26.19
Adaptability	Yes	53	25.49 *
	No	47	24.06 *
Interpersonal	Yes	53	27.04 *
	No	44	25.50 *
Communication	Yes	53	27.62 *
	No	45	25.67 *
Belief in teams	Yes	51	28.88
	No	47	28.47
Collective orient.	Yes	52	24.30
	No	46	24.63

Chapter 8

The ALL Practical Cognition[1] Framework

Robert J. Sternberg *Yale University, U.S.*
Elena L. Grigorenko *Yale University, U.S. and Moscow State University, Russia*
Guillermo Gil *Spanish Ministry of Education, U.S.*
Jennifer Hedlund *Yale University, U.S.*

1. This report draws for material on parts of Sternberg, R.J., Forsythe, G.B., Hedlund, J., Horvath, J.A., Wagner, R.K., Williams, W.M., Snook, S., and Grigorenko, E.L. (in press). *Practical intelligence in everyday life.* New York: Cambridge University Press.

Table of Contents

Table of Contents

Table of Contents

Summary

Practical cognition is what most people call common sense. It is the skill needed to adapt to, shape, and select everyday environments. Cognition as conventionally defined may be useful in everyday life, but practical cognition is indispensable. Without some measure of it, one cannot survive in a cultural milieu or even in the natural environment. In our work, we have studied many aspects of practical cognition, although we have concentrated on one particularly important aspect of it—*tacit knowledge*—the procedural knowledge one learns in one's everyday life that usually is not taught and often is not even verbalized. Tacit knowledge includes things like knowing what to say to whom, knowing when to say it, and knowing how to say it for maximum effect. In our work, we have studied tacit knowledge in populations as diverse as business managers, military leaders, university professors, elementary-school teachers, janitors, secretaries, salespeople, and U.S. and rural Kenyan children. Tacit knowledge is so-called because it usually starts off tacit, although over time it can come to be verbalized. It is measured by situational-judgment tests.

Our goal is not to denigrate the importance of more academically-based types of cognition, including literacy, numeracy, academic reasoning and so on. Rather, our argument is that more academic types of cognition are not enough—that successful prediction and, more importantly, understanding of performance in the everyday world requires assessment of practical as well as academic types of cognition.

1. Introduction

Practical (or everyday) cognition is different from the kind of cognition associated with academic success. There are any number of ways in which we see this difference in our everyday lives. We see people who succeed in school and who fail in work, or who fail in school but who succeed in work. We meet people with high academic-test scores who seem inept in their social interactions. And we meet people with low test scores who can get along effectively with practically anyone. Laypersons have long recognized a distinction between academic cognition (book smarts) and practical cognition (street smarts or common sense). This distinction is confirmed by research on the implicit theories of cognition held by both laypersons and researchers (Sternberg, 1985b; Sternberg et al., 1981).

1.1 Academic versus practical cognition

There may be any number of reasons for the apparent difference between academic and practical cognition. We argue that a major source of this difference is the sheer disparity in the nature of the kinds of problems one faces in academic versus practical situations. The problems faced in everyday life often have little relation to the knowledge or skills acquired through formal education or used in classroom activities. Consider the following example of an observation made by Richard Wagner of a garbage collector in Tallahassee, Florida.

> Tallahassee, priding itself on the service it provides to its citizens, requires garbage collectors to retrieve trash containers from the backyards of its residents. Each resident fills a large trash container in his or her backyard rather than placing standard-sized garbage cans on the curbside to be picked up. Trash collectors must locate and retrieve each full container from the backyard, heave it into the truck, and then drag the empty container back to each yard. Many of the garbage collectors are young high school dropouts who, because of their lack of education, might be expected to score poorly on cognition tests. On the surface, the job appears to be more physically than cognitively demanding. Each stop requires two trips to the backyard, one to retrieve the full can, and another to return it when it was empty.
>
> One summer it was noticed that the collection routine had changed after a new, older employee joined the crew. This change involved relaxing the constraint that each household retain the same container. Because the trash bins were issued by the city, and not purchased using personal funds, they were identical. The new routine consisted of wheeling the previous house's empty container to the current house's backyard, leaving it to replace the full can, which was in turn wheeled to the truck to be emptied. Once emptied, this can was wheeled to the backyard of the next house to replace its full can, and so on. The new routine required only one trip to each house, where the previous one required two trips. The new employee's insights cut the work nearly in half. This solution had eluded other garbage collectors and the managers who trained them.

Everyone encounters problems in which solutions are neither readily available nor readily derivable from acquired knowledge. This type of problem solving, frequently experienced in daily life, is referred to as *practical problem solving*. Such problems can be

experienced at the work place, or in school, the household, stores, movie theaters, or really, anywhere. There is no consensus on how to define practical problems encountered in life, but building on a distinction made by Neisser (1976), Sternberg and his colleagues (Sternberg, 1985a, 1997a; Wagner and Sternberg, 1986) have classified problems as academic or practical in nature. Academic problems tend to be (a) formulated by others, (b) well-defined, (c) complete in the information they provide, (d) characterized by having only one correct answer, (e) characterized by having only one method of obtaining the correct answer, (f) disembedded from ordinary experience, and (g) of little or no intrinsic interest.

Practical problems, in contrast to academic problems, tend to be (a) unformulated or in need of reformulation, (b) of personal interest, (c) lacking in information necessary for solution, (d) related to everyday experience, (e) poorly defined, (f) characterized by multiple "correct" solutions, each with liabilities as well as assets, and (g) characterized by multiple methods for picking a problem solution. Given the differences in the nature of academic and practical problems, it is no surprise that people who are adept at solving one kind of problem may well not be adept at solving problems of the other kind.

The cognitive skills that individuals exhibit in finding solutions to practical problems may be referred to as *practical cognitive skills* (Baltes, Dittman-Kohli, and Dixon, 1984; Berg, in press; Berg and Sternberg, 1985; Rogoff, 1982; Sternberg, 1985a, 1997a; Wagner, in press). When combined, these skills can be referred to as *practical cognition*, which is defined as cognitive skills that serve to find a more optimal fit between the individual and the demands of the individual's environment, whether by adapting to the environment, changing (or shaping) the environment, or selecting a different environment (Sternberg, 1985a; Sternberg, 1997a). The concept of practical cognition takes into account the distinction presented above between academic and practical tasks. The skills emphasized in formal schooling have limited value if they cannot be used to address practical, everyday problems.

1.1.1 Research on practical problem-solving skill

The research on practical cognition is becoming more and more central to mainstream psychology (see Berg and Klaczynski, 1996, for a review). Initially, the examination of practical cognition issued from a concern that the cognition of adults functioning largely outside the academic environment from the moment they obtained their academic degrees and virtually for the rest of their lives, was evaluated primarily by traditional tests of cognition constructed to predict academic success.

Various aspects of the meaning of the concept of practical cognition are expressed in a number of diverse constructs. Some researchers define everyday cognition as a specific expression of conventional skills that permit adaptive behavior within a distinct class of everyday-life situations (e.g., Willis and Schaie, 1986), whereas others stress the unique nature of practical skills (e.g., Neisser, 1976; Wagner, 1987). Most psychological studies of practical skills focus on solving problems that are ill-structured in their goals and solutions and are frequently encountered in daily life (at home, work, and in dealing with people) (e.g., Cornelius and Caspi, 1987; Denney, 1989).

A number of studies have addressed the relation between practical and academic cognition. These studies have been carried out in a wide range of settings, using a variety of tasks, and with diverse populations. We review some examples of research on problem solving and reasoning. For other reviews see Ceci and Roazzi (1994), Rogoff and Lave (1984), Scribner and Cole (1981), Sternberg and Wagner (1986, 1994), Voss, Perkins, and Segal (1991), and Wagner (in press). Taken together, these studies

show that skill measured in one setting (e.g., school) does not necessarily transfer to another setting (e.g., real-world task).

Several studies compared performance on mathematical types of problems across different contexts. Scribner (1984, 1986) studied the strategies used by milk processing plant workers to fill orders. Workers who assemble orders for cases of various quantities (e.g., gallons, quarts, or pints) and products (e.g., whole milk, two percent milk, or buttermilk) are called assemblers. Rather than employing typical mathematical algorithms learned in the classroom, Scribner found that experienced assemblers used complex strategies for combining partially filled cases in a manner that minimized the number of moves required to complete an order. Although the assemblers were the least educated workers in the plant, they were able to calculate in their heads quantities expressed in different base number systems, and they routinely outperformed the more highly educated white collar workers who substituted when assemblers were absent. Scribner found that the order-filling performance of the assemblers was unrelated to measures of school performance, including cognitive test scores, arithmetic test scores, and grades.

Another series of studies of everyday mathematics involved shoppers in California grocery stores who sought to buy at the cheapest cost when the same products were available in different-sized containers (Lave, Murtaugh, and de la Roche, 1984; Murtaugh, 1985). (These studies were performed before cost per unit quantity information was routinely posted). For example, oatmeal may come in two sizes, 10 ounces for $.98 for 24 ounces for $2.29. One might adopt the strategy of always buying the largest size, assuming that the larger size is always the most economical. However, the researchers (and savvy shoppers) learned that the larger size did not represent the least cost per unit quantity for about a third of the items purchased. The findings of these studies were that effective shoppers used mental shortcuts to get an easily obtained answer, accurate enough to determine which size to buy. A common strategy, for example, was mentally to change the size and price of an item to make it more comparable with the other size available. For example, one might mentally double the smaller size, thereby comparing 20 ounces at $1.96 versus 24 ounces at $2.29. The difference of 4 ounces for about 35 cents, or about 9 cents per ounce, seems to favor the 24-ounce size, given that the smaller size of 10 ounces for $.98 is about 10 cents per ounce. These mathematical shortcuts yield approximations that are as useful as the actual values of 9.80 and 9.33 cents per ounce for the smaller and larger sizes, respectively, and are much more easily computed in the absence of a calculator. When the shoppers were given a mental-arithmetic test, no relation was found between test performance and accuracy in picking the best values (Lave et al., 1984; Murtaugh, 1985).

Ceci and colleagues (Ceci and Liker,1986, 1988; see also Ceci and Ruiz, 1991) studied expert racetrack handicappers. Ceci and Liker (1986) found that expert handicappers used a highly complex algorithm for predicting post time odds that involved interactions among seven kinds of information. By applying the complex algorithm, handicappers adjusted times posted for each quarter mile on a previous outing by factors such as whether the horse was attempting to pass other horses, and if so, the speed of the other horses passed and where the attempted passes took place. By adjusting posted times for these factors, a better measure of a horse's speed is obtained. It could be argued that the use of complex interactions to predict a horse's speed would require considerable cognitive skill (at least as it is traditionally measured). However, Ceci and Liker reported that the successful use of these interactions by handicappers was unrelated to their overall cognitive ability.

A subsequent study attempted to relate performance at the racetrack to making stock-market predictions in which the same algorithm was involved. Ceci and Ruiz (1991) asked racetrack handicappers to solve a stock-market-prediction task that was structured similarly to the racetrack problem. After 611 trials on the stock-market task, the handicappers performed no better than chance, and there was no difference in performance as a function of overall cognitive ability. Ceci and Roazzi (1994) attribute this lack of transfer to the low correlation between performance on problems and their isomorphs. "Problem isomorphs" refer to two or more problems that involve the same cognitive processes but that use different terminology or take place in different contexts.

The same principle that applies to adults appears also to apply to children. Carraher, Carraher, and Schliemann (1985) studied Brazilian children who, for economic reasons, often worked as street vendors (see also Nuñes, 1994). Most of these children had very little formal schooling. Carraher et al. compared the performance of these children on mathematical problems that were embedded in a real-life situation (i.e., vending) to problems presented in an academic context (e.g., 2 + 4 = ?). The children correctly solved significantly more questions that related to vending than they did math problems that were academic in nature. When the academic problems were presented as word problems (e.g., If an orange costs 76 cruzeiros and a passion fruit cost 50, how much do the two cost together?), the rate of correct responses was substantially better, but still not as high as when the problems were presented in the context of vending.

This lack of transfer also appears to work in the reverse direction. For example, Perret-Clermont (1980) found that many school children had no problem solving paper-and-pencil arithmetic questions, but could not solve the same type of problem in a different context (e.g., counting bunches of flowers). That is, school children may fail to transfer the academic knowledge to everyday problems.

Roazzi (1987) found similar results when comparing street-vendor children to middle-class school children. He compared the performance of children on a class-inclusion task. To assess the performance of the street-vendor children, the researcher posed as a customer and asked questions about the items to find out if the children understood the relationship among classes and subclasses of food (e.g., mint and strawberry chewing gum as part of the class "chewing gum"). At a later time the same children were given a formal test with the same logical structure, but that was irrelevant to their street-vending jobs. The middle-class children were given the same two tests. Street-vendor children performed significantly better on the class-inclusion task in the natural than in the formal context, whereas middle-class children were more successful on the formal version of the task.

Additional research has shown that the use of complex reasoning strategies does not necessarily correlate with overall cognitive ability. Dörner and colleagues (Dörner and Kreuzig, 1983; Dörner, Kreuzig, Reither, and Staudel, 1983) studied individuals who were asked to play the role of city managers for the computer-simulated city of Lohhausen. A variety of problems were presented to these individuals, such as how best to raise revenue to build roads. The simulation involved more than one thousand variables. Performance was quantified in terms of a hierarchy of strategies, ranging from the simplest (trial and error) to the most complex (hypothesis testing with multiple feedback loops). No relation was found between overall cognitive ability and complexity of strategies used. A second problem was created to cross-validate these results. This problem, called the Sahara problem, required participants to determine the number of camels that could be kept alive by a small oasis. Once again, no relation was found between overall cognitive ability and complexity of strategies employed.

The above studies indicate that demonstrated skills do not necessarily correspond between everyday tasks (e.g., price-comparison shopping) and traditional academic tasks (e.g., math achievement tests). In other words, some people are able to solve concrete, ill-defined problems better than well-defined, abstract problems that have little relevance to their personal lives, and vice versa. Few of these researchers would claim, however, that academic skills are totally irrelevant to performance in these various contexts. There is evidence that conventional tests of cognition predict both school performance and job performance (Barrett and Depinet, 1991; Schmidt and Hunter, 1998; Wigdor and Garner, 1982). What these studies do suggest is that there are other aspects of cognition that may be independent of academic cognition and that are important to performance, but that largely have been neglected in the measurement of cognition. We also observe this incongruity between conventional notions of real-world skills in research on age-related changes in cognitive skill.

1.1.2 The fun of growing older: Do age-related patterns in practical cognition resemble those in conventional cognition?

Throughout the century of existence of cognitive psychology, many cognitive variables (mostly those contributing to the *g*-factor—for review, see Berg, in press; Sternberg and Berg, 1992) have been found to be associated with age across the life-span. Most of these associations are rather complex and of curvilinear nature, reflecting rapid growth during the years of formal schooling and slow decline thereafter (Salthouse, 1998). However, the results of research also suggest somewhat different developmental functions for changes in performance on various kinds of cognitive tasks across the adult life span. In particular, data show that older adults commonly report growth in practical skills over the years, even though their academic skills decline (Williams, Denney, and Schadler, 1983).

As for specific cognitive functions, cognition during adulthood is characterized, on one hand, by losses in the speed of mental processes, abstract reasoning, and specific characteristics of memory performance (see Salthouse, 1991, for a review) and, on the other hand, by gains in the metacognitive skill to integrate cognitive, interpersonal, and emotional thinking in a synthetic understanding of the world, self, and others (Labouvie-Vief, 1992, for a review).

The most commonly used theoretical framework adapted for the interpretation of findings on age-related changes in cognitive performance is that of fluid and crystallized cognitive skills (Horn, 1994; Horn and Cattell, 1966). *Fluid* skills are those required to deal with novelty, such as in the immediate testing situation (e.g., discovering the pattern in a figure sequence). *Crystallized* skills are represented by accumulated knowledge (e.g., finding a synonym of a low-frequency word). Utilizing this distinction, many studies have demonstrated that fluid skills are relatively susceptible to age-related decline, whereas crystallized skills are relatively resistant to aging (Dixon and Baltes, 1986; Horn, 1982; Labouvie-Vief, 1982; Schaie, 1977/1978), except near the end of one's life.

In addition, Willis and Schaie (1986) studied the relationships between fluid and crystallized skills and everyday cognition (the latter being defined as the skill to perform core activities of independent life—e.g., cooking, managing finances, or using the telephone and measured by a variant of the ETS Basic Skills Test) in the elderly. The researchers reported substantial correlations between performance on the Basic Skills Test and a measure of fluid (r = .83) and crystallized (r = .78) skills.

The majority of these findings, however, were obtained in the framework of cross-sectional methodologies, that is, by comparing different groups of individuals of various ages. When the same individuals are followed across time in the framework of longitudinal

design, the findings show that, with respect to fluid cognition, decline does not generally begin until the sixties and loss of crystallized cognition occurs almost a decade later, in the seventies (Schaie, 1996).

In addition, even when there are age-based group differences in cognitive performance, there is extensive interindividual variability for specific cognitive skills within age groups. For instance, Schaie (1996), although consistently reporting mean cross-sectional differences in overall cognitive performance, pointed out impressive variability within age groups. To quantify this variability, Schaie (1988) investigated the overlap in distributions of cognitive performance among young adults and the elderly. Even in the group of eighty and over the overlap was about 53 percent, scoring well above the mean of their age group. In other words, half or more than half of individuals in the late age groups perform comparably to a group of young adults on measures of both crystallized and fluid cognition.

Moreover, there is also a considerable amount of interindividual variability in the longitudinal patterns of decline, maintenance, and improvement. Specifically, Schaie and Willis (1986) categorized older individuals (the group mean age was 72) into those who decline and those who remained stable in their performances on the Primary Mental Abilities Test (using the space and reasoning subtests) over a period of fourteen years. Forty-seven percent of the sample remained stable on both measures, whereas only 21 percent declined on both measures. Some of these individuals were followed into their eighties, and virtually none of them showed universal descent across all five subtests of the Primary Mental Abilities Test (Schaie, 1989). It is thought that those who show age-related maintenance and improvement in cognitive development differ from those showing decline on a constellation of factors, including educational background, occupational pursuits, health history, life habits, and such personality styles as rigidity and flexibility (Schaie, 1996).

The trend of cognitive development across the lifespan, however, appears to be yet somewhat different for practical skills. Williams et al. (1983) interviewed men and women over the age of 65. The questions posed to these adults had to do with their perception of age-related changes in their skill to think, reason, and solve problems. Surprisingly enough, the responses obtained from these adults were largely contradictory to the view that late development of cognition consists of decline (see Berg, in press, for review). In the Williams et al. study (1983), 76% of the elderly adults believed that their skill to think, reason, and solve problems had actually increased over the years, with 20% reporting no change and only 4% reporting that their skills had declined with age. The researchers confronted the participants with the overwhelming evidence of decline in conventional test performance upon completion of formal schooling, but the explanation of the elderly people was that they were talking about solving kinds of problems different from those found on psychometric tests. The problems they had in mind when answering the interviewer's questions were those of an everyday or financial nature. Of course, these responses might be simply discounted as self-deceiving and self-reassuring, but a number of formal psychological studies within the last decade have provided significant support for the claim made by the elderly in the Williams et al. (1983) study.

In particular, the idea that practical and academic skills might have different developmental trajectories was supported in a number of studies (see Berg and Klaczynski, 1996, for a review). Denney and Palmer (1981) were one of the first research teams to demonstrate this discrepancy. They compared the performance of adults (aged 20 through 79) on traditional analytical reasoning problems (e.g., a "twenty questions" task) and a problem-solving task involving real-life situations (e.g., "If you were traveling by car and got stranded out on an interstate highway during a blizzard, what would you

do?"). One of the many interesting results obtained in this study was a difference in the shape of the developmental function for performance on the two types of problems. Performance on the *traditional* problem-solving task or cognitive measure declined almost linearly from age 20, onward. Performance on the *practical* problem-solving task increased to a peak in the 40- and 50-year-old groups, declining thereafter. Expanding on this line of research, Smith and colleagues (Smith, Staudinger, and Baltes, 1994) compared responses to life-planning dilemmas in a group of younger (mean age 32) and older (mean age 70) adults. Unlike the results of studies of aging and academic skills, which demonstrated the superior performance of younger adults over the elderly, in this study, young and older adults did not differ. In addition, each age-cohort group received the highest ratings when responding to a dilemma matched to their own life phase.

Similar results were obtained in a study by Cornelius and Caspi (1987). They studied adults between the ages of 20 and 78. These researchers examined relationships between performance on tasks measuring fluid cognition (letter series), crystallized cognition (verbal meanings), and everyday problem solving (e.g., dealing with a landlord who won't make repairs, filling out a complicated form, responding to criticism from a parent or child). Performance on the measure of fluid skill increased from age 20 to 30, remained stable from age 30 to 50, and then declined. Performance on the everyday problem-solving task and the measures of crystallized skill increased through age 70.

Likewise, the neofunctionalist position, advanced by Baltes and his associates (Baltes, 1987; Baltes et al., 1984; Baltes, Smith, and Staudinger, 1992; Dittmann-Kohli and Baltes, 1990) acknowledges that, although some aspects of cognitive functioning estimated via traditional tests may decline with age, stability and growth also exist, if to a lesser extent. The approach of Baltes and his colleagues also operates within the constructs of fluid and crystallized cognition, although a different emphasis is placed on the relative roles and meanings of these two kinds of cognition. Here, both aspects of cognition are considered as coequals in defining the developmental course of cognition. In general, Baltes argues that crystallized cognition has been too narrowly defined, and that its importance increases as one moves into adulthood and old age. In this sense, it may be inappropriate to associate a decrease in fluid cognition with an average decline in cognitive competence. Baltes and his associates see adult cognitive competence in terms of a dual-process model. The first process, called the *mechanics* of cognition, is concerned with developmental change in basic information processing that is genetically driven and assumed to be knowledge-free. With aging, there is a biologically-based reduction in reserve capacity (Baltes, 1987; Baltes et al., 1992). The second process, *pragmatic* cognition, relates the basic cognitive skills and resources of the first process to everyday cognitive performance and adaptation. Measures of pragmatic cognition within select domains are viewed as tapping skills more characteristic of adult cognitive life than are traditional psychometric measures of cognitive skills. Similar to empirical findings on the distinction between fluid and crystallized cognition, Baltes, Sowarka, and Kliegl (1989) showed that the mechanics of cognition tend to decline with age almost linearly, whereas the pragmatics of cognition tend to maintain relative stability throughout adulthood. For example, whereas linear declines were found in the speed of comparing information in short-term memory (i.e., aspects of cognitive mechanics), no age differences were registered for measures of reasoning about life planning (i.e., aspects of cognitive pragmatics). Cognitive skills are assumed to operate on content domains involving factual and procedural knowledge; they are regulated by higher-level, trans-situational, procedural skills and by higher-order reflective thinking (metacognition), all of which define the "action space" in which problem solving occurs within a given individual. According to this approach, successful aging entails limiting one's tasks and avoiding excessive demands. Baltes and Baltes (1990) use the concept of selection to refer to a self-imposed restriction in one's life to fewer domains of functioning

as a means to adapt to age-related losses. It is assumed that by concentrating upon high-priority domains and devising new operational strategies, individuals can optimize their general reserves (Baltes, 1993). By relating adult cognition to successful cognitive performance in one's environment, this position acknowledges that not all tasks are equally relevant for measuring cognition at different ages (Baltes et al., 1984; Baltes et al., 1992).

Specific manifestations of pragmatic cognition are said to differ from person to person as people proceed through selection, optimization, or compensation (Dittmann-Kohli and Baltes, 1990). Selection refers simply to diminishing the scope of one's activities to things that one is still able to accomplish well, despite a diminution in reserve capacity. Thus, research shows that elderly people tend to leave jobs that require quick sensorimotor responses (Barrett, Mihal, Panek, Sterns, and Alexander, 1977). Optimization refers to the fact that older people can maintain high levels of performance in some domains by practice, greater effort, and the development of new bodies of knowledge. Compensation comes into play when one requires a level of capacity beyond remaining performance potential. For example, Salthouse (1984) was able to show that older typists, although slower on several simple speeded reaction-time tasks, were able to compensate for this deficit and maintain their speed by reading further ahead in the text and planning ahead. According to Salthouse and Somberg (1982), age-related decrements at the "molecular" level (e.g., in speed of execution of the elementary components of typing skill) produce no observable effects at the "molar" level (i.e., the speed and accuracy with which work is completed).

Charness (1981) showed similar effects with older chess players, who exhibited poorer recall in general, but were better able to plan ahead than younger, less experienced players. In related studies, older adults have been found to compensate for declines in memory by relying more on external memory aids than do younger adults (Loewen, Shaw, and Craik, 1990). Older adults must often transfer the emphasis of a particular task to skills that have not declined in order to compensate for those that have (see Bäckman and Dixon, 1992, for a review of these issues). In other words, when a task depends heavily on knowledge, and speed of processing is not a significant constraint, peak performance may not be constrained in early-to-middle adulthood (Charness and Bieman-Copland, 1994). As an example, consider chess competitions by correspondence. In these "chess-by-mail" competitions, players are permitted three days to deliberate each move. The mean age of the first-time winners of one postal world championship is 46 years old. In contrast, the peak age for tournament chess, where deliberation averages three minutes per move, is about 30 years old (Charness and Bosman, 1995).

A series of studies on the relationship between aging and cognitive efficiency in skilled performers attested to the compensatory and stabilizing role of practical cognition (Baltes and Smith, 1990; Charness and Bosman, 1990; Colonia-Willner, 1998; Hartley, 1989; Willis, 1989). Sternberg and colleagues' studies of tacit knowledge in the domains of business management, sales, and academic psychology showed increases in tacit knowledge with age and experience across groups of undergraduates, graduate students, and professionals (Sternberg, Wagner, Okagaki, 1993; Wagner, 1987; Wagner, Rashotte, and Sternberg, 1994; Wagner and Sternberg, 1985). Colonia-Willner (1998) found evidence that older managers who performed at the highest levels on average had high levels of tacit knowledge—even though on average they had relatively low scores on psychometric reasoning measures. In addition, Colonia-Willner pointed out an interesting detail: even though tacit knowledge of managerial skills was shown to be related to some indicators of job success for the total sample of bank managers, the relative weight of this knowledge was higher for the highest success group (that group rewarded most highly). It might be that job-related tacit knowledge is especially

important for detecting super-achievers among a fairly restricted, high-achieving, conventional population of managers engaged in heterogeneous activities.

Moreover, a series of training studies, conducted in Germany (Baltes et al., 1984; Baltes et al., 1992) and the U.S. (Schaie, 1986; Schaie and Willis, 1986; Willis and Schaie, 1994), have shown that older individuals still have a great deal of potential plasticity, or reserve capacity for development. The results demonstrated that intervention can lead to significant gains in skills such as problem-solving tasks (Denney, 1979), perceptual speed (Hoyer, Labouvie, and Baltes, 1973), and fluid cognition (Baltes and Lindenberger, 1988; Willis, 1987). Intervention research generally targeted those skills which have been shown to decline the most (i.e., fluid cognition and processes representative of the mechanisms of cognition).

In general, results from intervention studies convincingly demonstrated the remarkable plasticity of human cognition in the elderly (see Willis, 1987 for a review). In the German studies, better performance was demonstrated for (1) target training (Baltes and Willis, 1982; Willis, Blieszner, and Baltes, 1981), (2) independent self-practice (Baltes et al., 1989; Hayslip, 1989a, 1989b), and (3) removed time constraints (Hofland, Willis, and Baltes, 1981). Willis and Schaie, 1986; Schaie and Willis, 1986; Willis and Schaie, 1994) obtained similar findings within a longitudinal design.

These results were replicated in a second follow-up study conducted in 1991 with both new participants and participants from the original training study. Specifically, results from the Seattle Training Study, a component of the Seattle Longitudinal Study (Schaie, 1996) indicated that the performance of the elderly can be successfully impacted in such a way that older adults' performance is boosted back to the level at which they performed more than a decade before. The Seattle researchers set up five one-hour sessions aimed at training the elderly adults' spatial and reasoning skills. The training had differential impact on certain subgroups of the elderly population. For those who had shown decline on either of the Primary Mental Skill Test subtests over the preceding fourteen-year period, training was effective in returning their performance nearly to the original level. For those who had remained stable over the preceding fourteen-year period, training raised their performance beyond the level they performed at fourteen years prior to the training. In addition, the training has been found to be effective, not only in the short run, but over seven years (Neely and Backman, 1993; Willis and Nesselroade, 1990).

One of the outcomes of these studies is the realization that longer and more structured training seems to be necessary for remediation in the very old (Schaie, 1994; Willis, 1989). The importance of these studies is that they suggest that cognitive decline in many individuals may be due to disuse of certain cognitive skills, and that remediation is possible for a significant number of participants, especially for the young-old (Schaie, 1994; Willis, 1990; Willis and Schaie, 1994).

The developmental trajectory of everyday cognition has been examined by a number of researchers (see Berg, in press; Berg and Klaczynski, 1996, for review). The summary of the field today is that the pattern of age differences in everyday cognition differs dramatically depending on how problems to be solved are defined and what criteria are used for optimal problem solving. For example, Berg, Klaczynski, Calderone, and Strough (1994), studying participants' own ratings of how effective they were in solving their own everyday problems, did not find any age differences. Denny and her colleagues (Denney and Palmer, 1981; Denney and Pearce, 1989) utilized the number of "safe and effective solutions" as the criterion for optimal problem solving and found that the highest number of such solutions was generated by middle-aged adults, with both younger and older adults offering fewer solutions. Cornelius and Caspi (1987), using the closeness between participants' ratings of strategy effectiveness and a "prototype"

of the optimal everyday problem solver as the criteria, found an increase in everyday problem-solving skill with adult age.

A number of studies have examined everyday problem solving with a neo-Piagetian approach to cognitive development in adulthood (Labouvie-Vief, 1992). According to this paradigm, in middle and late adulthood, the formal operational reasoning of late adolescents and young adults, with its focus on logic, is replaced by more sophisticated mental structures distinguished by relativistic reasoning based on synthesizing the irrational, emotive, and personal. Specifically, Blanchard-Fields (1986, 1994; Blanchard-Fields, and Norris, 1994) stated that, when dealing with social dilemmas, older adults are superior to younger adults in their integrative attributional reasoning (i.e., reasoning based on the integration of dispositional and situational components).

To conclude, there is reason to believe that the developmental trajectories of skills utilized to solve strictly academic problems do not coincide with the trajectories of skills used to solve problems of a practical nature.

1.2 What develops in practical cognition?

The evidence supporting the supposition that practical cognition has a different developmental trajectory than academic cognition supports the etiological independence (not necessarily complete) of practical and academic skills but is only one of many research advances revealing the developmental mechanisms of practical cognition. Developmental research on practical skills is still in its early stages. However, data available at this point shed some light on what Sinnott (1989) called the *chaotically complex* reality of practical problem solving; evidence supports the existence of different developmental trajectories (maintenance, improvement, and decline) across the life span without a pronounced preference for any single one.

There is no formal theory of the stages of the development of practical cognition (Berg, 1994). Some results, however, suggest that the difference in performance on practical and analytical tasks is observed rather early. Freeman, Lewis, and Doherty (1991) have shown that the performance of preschoolers on the false-belief tasks (e.g., tasks involving the formation of false beliefs and expecting children to determine and overcome their false nature) is better if they are asked to act out answers rather than to give them verbally. The researchers suggest that the reason for this discrepancy is that early implementation of a theory of intentionality is "only" practical. In other words, preschool children are able to distinguish between true and false expectations and true and false causes, but do it by carrying out practical actions (e.g., acting with the right object) rather than by explaining why those particular objects should be chosen. These and other findings contribute to the hypothesis that reflective awareness and verbalization emerge gradually from the implicit practical cognition organizations which are their necessary precursors (e.g., Bickhard, 1978; Karmiloff-Smith, 1988).

Developmental research on practical cognition is moving in a number of directions, each of which might help us to detect the internal mechanisms of its development. Most of the work is centered on specific characteristics of practical tasks. The assumption here is that if we understand the differences in the ways these tasks are formulated and solved at different stages of development, we will be closer to understanding the developmental dynamics of practical cognition. Drawing on the distinction made earlier between academic and practical tasks suggests five main directions of research: (1) studies of developmentally variable contexts of practical problem solving; (2) studies of developmental changes in the content of practical problems encountered at different stages of development; (3) studies of the developmental diversity of the goals of practical problem solving; (4) studies of differential strategies utilized in practical problem solving

at different periods of development; and (5) studies on developmental variation in problem interpretation and definition.

1.2.1 Context of practical problem solving

There is virtually unanimous agreement on the centrality of context for understanding practical problem solving. This view, which holds that practical problem solving cannot be separated from the context in which it unfolds, is referred to as the contextual perspective (e.g., Dixon, 1994; Wertsch and Kanner, 1994). In general, the metaphor used to describe the contextual approach is that of trying to follow forever changing events (i.e., the life course is represented as being a series of changing events, activities, and contexts). When applied to studies of practical problem solving, this perspective assumes that (1) the demands posed by these contexts vary across development; (2) strategies accomplishing adaptation differ across contexts; (3) these strategies also differ across individuals; and, finally, (4) the effectiveness of everyday problem solving is determined by the interaction of individual and context (Berg and Calderone, 1994). Several studies have found that the context in which the problem occurs (e.g., family, work, or school) impacts everyday problem solving in all its components (content, goal, and strategy).

Consider the following examples. Ceci and Bronfenbrenner (1985; Ceci, 1990), employing a dual context paradigm, have conducted a series of studies concerning the impact of physical and social contexts on cognition. The dual context paradigm proposes that children be made to perform the same task in two or more contexts. The assumption here is that some settings elicit more effective forms of cognition than do others by stimulating or activating different strategies. The Ceci-Bronfenbrenner view is that a task perceived in a modified form might recruit a set of strategies acquired previously but not elicited by the original, unmodified task. (For example, a video-game task, which is a modification of a simple task requiring a participant to follow the movements of dots, might recruit strategies that the dot task alone would not.) Cohen (1996) studied the mathematically-oriented activity of 3- and 4-year olds and found that, when mathematical operations were embedded in the broader context of a "play-store" setting, children were able to solve problems that exceeded an age-appropriate level of difficulty. In addition, the children satisfied the demands of the task in using a variety of solution strategies.

One of the most interesting developments in studies on context and practical problem solving concerns the effect of compensation: the phenomenon in which gains in (mostly) practical cognition balance out age-related decrements in others. Researchers argue that compensation—considered in terms of the dynamic relationship between the individual's changing cognitive skills and expectations of performance, on the one hand, and shifting contextual demands, on the other hand—should be viewed as central to cognitive aging (e.g., Dixon, 1994). One example of practical cognition compensating for declines in *g*-based cognitive performance is older adults' effective use of external aids. One common source of external cognitive aid is other people. For example, Dixon and his colleagues (Dixon, 1994) explored the extent to which older and younger adults use same-age collaborators in solving memory problems and found that older adults use previously unknown collaborators to boost their performance levels to a much greater extent than do younger adults.

Two other important characteristics of the context in which practical problem solving occurs, which might explain some aspects of the observed development variability in practical cognition, are the complexity and familiarity of the context.

As for the complexity of the environment in which practical cognition unfolds, one variable that has been pointed out as extremely important for shaping the development of practical skills in adulthood is that of the immediate conditions and demands of work (see Schooler, in press, for a review). For example, Kohn and Schooler (1983), in a group of men between the ages of 24 to 64, longitudinally studied the link between the extent to which one's work-related activities involve independent thought and judgment and workers' flexibility in dealing with complex cognitive demands. They found that the more the substantive complexity of one's job, the greater the incremental gains in cognitive performance over a ten-year period. Even more astounding, a similar relationship between job complexity and cognitive performance was revealed for women doing complex housework (Schooler, 1984). Moreover, K.A. Miller and Kohn (1983) found that individuals with higher flexibility in dealing with complex cognitive activities tended to engage in more stimulating and demanding cognitive activities (e.g., reading books versus watching television). The major criticism of this nonexperimental evidence of the cognitive effects of doing complex work (whether in the work place or the household) is that these designs are unable to rule out the possibility that individuals who maintain their cognitive functioning are more capable of pursuing and staying in challenging work environments. Yet, even though the causal path is difficult to infer among individuals, the evidence that among individuals more cognitively complex work leads to enriched cognitive functioning deserves attention and more thorough investigation.

Regarding familiarity or experience with the domain in which practical problem solving is carried out, studies have demonstrated that cognitive performance is greater for both young and older adults when individuals are given either familiar materials (Smith and Baltes, 1990) or a chance to practice prior to assessment (Berg, Hertzog, and Hunt, 1982). Yet, results are ambiguous as to whether differential familiarity is a factor that can help to explain age differences in practical problem solving (Denney and Pearce, 1989).

Researchers reported, for example, that older adults perceived traditional cognition tests as less familiar than did young adults (Cornelius, 1984). Therefore, when younger and older adults are compared on conventional cognition tests, older adults might look worse because these tests are less familiar to them and they may have forgotten how to evoke specific strategies relevant to situations of cognitive assessment.

To explore the importance of the familiarity factor, several studies have been carried out in which younger and older adults were asked to solve problems that were constructed to be more familiar or more normative for one age group or the other. For example, Denney and colleagues (Denney, Pearce, and Palmer, 1982) showed that, in adults, the more normative for their age group everyday problems are, the better their performance is. Similarly, Smith and Baltes (1990) found that adults perform best when the problems are more normative for their age group. As Berg (in press) pointed out, memory research utilizing the usage of tasks with familiar materials (e.g., remembering words that were in frequent use during their adulthood years versus contemporary equivalents) is consistent in showing that older adults tend to perform better with materials more familiar to them (Barret and Watkins, 1986; Worden and Sherman-Brown, 1983).

1.2.2 Content of practical problem solving

The main hypothesis underlying this line of research is that the content of practical problem solving differs at different stages of development. The literature published to verify this hypothesis contains heterogeneous evidence; some is supportive (e.g., Aldwin,

Sutton, Chiara, and Spiro, 1996) and some is not supportive (e.g., Folkman, Lazarus, Pimley, and Novacek, 1987) of the assertion that individuals of different ages experience different everyday problems.

Berg and colleagues (Berg and Calderone, 1994; Sansone and Berg, 1993) asked preschoolers, teenagers, college students, and older adults to describe a recent problem (hassle, conflict, challenge, and so on) that they had experienced and to describe the problem in as much detail as possible. The intent was to investigate whether the types of domains of problems remain constant across development or whether different types of problems would appear for different age groups. The researchers found significant variation in the content of everyday problems across development. The everyday problem-solving content for 5-6-year-olds consisted predominantly of problems dealing with family (e.g., disagreements with family members) and assigned responsibilities (e.g., home chores). For 11 to 12-year-olds, everyday life problems centered on school and after-school activities and environments. No single content area dominated the everyday life of college students, and their salient problems had to do with free time, work, friends, family, and romantic relationships. Finally, the everyday problem solving of the older adults centered on the family context and health.

Barker (1978) suggested that the content of practical problem solving is determined by the ecological characteristics of a given developmental period. They carried out detailed observations of settings inhabited and experienced by elementary school children on a daily basis and found that children most frequently occupy settings embedded in schooling and family life. This piece of work is unique in terms of its thoroughness and attention to details; however, based on sporadic evidence accumulated in research on developmental life tasks, the general assumption in the field is that the content of the practical problem solving of adults differs in a variety of ways across the life span. In other words, it might be impossible to carry out Baker et al.-like studies in all ecological settings encountered in adulthood, but it might be possible to target the few that appear to be crucial at specific developmental periods. Specifically, it has been shown that (1) college students' tasks are primarily aimed at succeeding academically, forming social networks, developing an identity, and separating from family (Cantor, Norem, Neidenthal, Langston, and Brower, 1987); (2) adults focus on a variety of tasks, ranging from starting a family and a career in young adulthood, through the pragmatic tasks of middle adulthood, to adapting to impairments of health and adjusting to retirement during old and advanced old age (Baltes et al., 1984; Havinghurst, 1972; Neugarten Moore, and Lowe, 1968).

1.2.3 Goals of practical problem solving

The goal-directedness (e.g., Goodnow, 1986; Scribner, 1986; Wertsch, 1985) of practical problem solving is one of the most often cited characteristics of practical cognition in application. Therefore, the second line of research concerns the developmental trajectories of goals of practical problem solving.

Strough, Berg, and Sansone (1996) showed that there is developmental variation in the types of goals underlying everyday problem solving. The profile of this developmental variation reflects developmental life tasks (Cantor, 1990). Specifically, preadolescents reported more goals for task improvement, and a large portion of their problems involved the school context. Interpersonal goals appeared to be more salient to middle-aged adults than to preadolescents. Preadolescents, however, reported more other-focused assistance-recruiting goals than did adults. Older and middle-aged adults reported more physical goals than did younger individuals, and the adult group as a whole reported more affective goals than did preadolescents.

Klaczynski, Laipple, and Jurden (1992) studied practical cognition among adolescents in college-preparatory or vocational-training tracks. Depending on the chosen developmental life-track, adolescents in the two groups differed in their interpretation of practical problem situations. In particular, vocational students were concerned primarily with goals involving the acquisition of adult status, such as marriage, steady employment, and independence. College-preparatory students, on the other hand, reported more achievement-oriented goals, such as doing well in school, gaining admission to quality colleges, and scoring well on entrance exams.

Belief in the plasticity and fluidity of human developmental goals throughout the life span is also reflected by the notion that there is no single outcome or endpoint to cognitive development in general, or to the development of practical cognition in particular (e.g., Rogoff, 1982). The implication of this line of reasoning is that the individual and his or her context form a complex systemic unit; changes in the unit shape the content, dynamics, and adaptability of the individual's cognitive functioning in specific contexts. Thus, there is no "ideal" trajectory of cognitive development, and there is no optimal instrument assessing cognitive functioning equally well at all periods of the life span.

1.2.4 Practical problem-solving strategies

One of the main research trajectories in the field of practical cognition focuses on strategies utilized in problem solving. Among the central characteristics of strategies discussed in the research literature of the past 20 years (Belmont and Butterfield, 1969; Berg, 1989; Brown, 1975; Flavell, 1970; Naus and Ornstein, 1983; Pressley, Forest-Pressley, Faust and Miller, 1985) are selectivity, goal-directedness, and intentionality. Many developmental researchers have been especially interested in strategy selection as both an individual and a developmental indicator of everyday problem-solving performance (e.g., Frederiksen, 1986; Frederiksen, Jensen, and Beaton, 1972; Lazarus and Folkman, 1984).

Most of the early developmental work on everyday problem solving has been carried out under the assumption that individuals' chosen strategies can be compared irrespective of the developmental variation in the goals motivating these strategies (Band and Weisz, 1988; Berg, 1989; Cornelius and Caspi, 1987; Folkman et al., 1987). The major theoretical hypothesis dominating the field is that greater experience with everyday problems leads to better problem solving (Baltes et al., 1984, Denney, 1982). This claim assumes that a particular type of strategy—e.g., primary control reflected in independent coping and problem-focused action—is a more effective way of dealing with various problems than is some other strategy—e.g., secondary control reflected in reliance on others and emotion-focused action (Denney, 1989; Folkman et al., 1987). For example, self-action was the strategy most frequently mentioned across all ages in a study of reported everyday problems (Berg, Strough, Calderone, Sansone, and Weir, 1998). Problem-focused action was most frequently mentioned for hypothetical problems (Blanchard-Fields, Jahnke, and Camp, 1995). Developmental differences have been encountered, suggesting that secondary control strategies, emotion-focused strategies, and dependence on others increases across early childhood (Band and Weisz, 1988), with further elevation in later adulthood (Brandtstaedter and Greve, 1994; Denney and Palmer, 1981; Folkman et al., 1987; Heckhausen and Schultz, 1995). For instance, researchers (Band and Weisz, 1988) found that older children were more likely to use secondary control strategies, such as efforts to modify the subjective psychological state of the self to better suit the present conditions of the problem, whereas younger children were more likely to use primary control strategies, such as efforts to influence the problem so that it meets the problem solver's expectations.

The empirical literature, however, does not uniformly support the claim that "more experience equals better problem solving" (Baltes, 1997; Berg, 1989; Cornelius and Caspi, 1987). Recent research suggests that strategies are differentially effective depending on the context of the everyday problem (Berg, 1989; Ceci and Bronfenbrenner, 1985; Cornelius and Caspi, 1987; Scribner, 1986). Thus, Cornelius and Caspi (1987) showed that different types of strategies (problem-focused action, cognitive problem analysis, passive-dependent behavior, and avoidant thinking and denial) were viewed as differentially effective in different contexts.

Findings regarding the localization of age differences are also somewhat contradictory. The often-cited trend in the literature is that older adults tend to use more secondary control (e.g., Heckhausen and Schulz, 1995) and less problem-focused action or primary control (Folkman et al., 1987) when compared to younger adults. Blanchard-Fields et al. (1995) found minimal age differences in problem-focused action. Furthermore, Berg et al. (1998) reported age differences for older adults only, with older people using relatively less cognitive regulation and more self-action than either college students or middle-aged adults. The situation has become even less transparent, with Aldwin et al. (1996) showing that, for the most part, age differences existed among adults only when individuals' strategies were assessed through a checklist; these distinctions were greatly reduced when individuals' strategies were elicited through open-ended interviews.

One of the possible explanations for the heterogeneity of these findings is that what develops over time is sensitivity to specific contexts. In other words, the repertoire of dealing with everyday problems is rather broad, and different modules of problem solving are used in different situations; in many ways, consistency across situations may be maladaptive (Mischel, 1984). Some researchers argue that successful everyday problem solving will involve carefully fitting strategies to the specific demands of a problem and modifying these strategies in response to changes in the problem (Berg and Sternberg, 1985; Rogoff, Gauvain, and Gardner, 1987; Scribner, 1986). And sensitivity to the contextual features of a problem is characteristic of a developmental factor (Mischel, 1984; Rogoff et al., 1987). Others, on the contrary, suggest that these strategies become less context-dependent with age (e.g., Kreitler and Kreitler, 1987).

Yet another, although not contradictory possibility, is that the lesson derived from experience with everyday problems is how to avoid getting into everyday problems (Berg, 1989). Thus, it is plausible that no simple relation between kind of experience and everyday problem-solving skill is likely to exist. Moreover, researchers have presented evidence demonstrating that so-called effective-across-all-contexts (e.g., primary) strategies fail in situations in which so-called ineffective strategies (e.g., relinquishing) work (Berg, Calderone, and Gunderson, 1990, as cited in Berg and Calderone, 1994). Certain kinds of experience may be differentially related to success at solving particular kinds of everyday problems, and development might better be construed as individuals becoming increasingly capable of modifying their strategies or avoiding potentially problematic situations (Berg, 1989; Rogoff et al., 1986).

Another line of research focuses on studying individual differences that appear to lead to more optimal problem-solving performance (e.g., Ceci and Liker, 1986; Denney, 1989; Willis and Schaie, 1986). Many factors (e.g., conventional cognitive skills, personality traits, social skills, achievement motivation) have been shown to impact the utilization of strategies in everyday problem solving (e.g., Ceci and Liker, 1986; Charness, 1981; Kuhn, Pennington, and Leadbeater, 1983), but no specific constellations of these factors were found to be better predictors of effective problem solving.

1.2.5 Problem interpretation (Definition)

In an attempt to systematize the literature on the development of everyday problem solving, Berg and colleagues have introduced the concept of "problem interpretation" (Berg and Calderone, 1994; Sansone and Berg, 1993) or "problem definition" (Berg et al., 1998). The problem interpretation arises at the edge of the context and the individual and, in essence, is the transaction of the individual with his or her context. The problem interpretation derives from features of both the individual and the context, but it might selectively engage all or only some features. Berg and her colleagues argue that such individual and contextual features may have different weights and may be differentially combined at different stages of development; thus, the search for developmental variation in everyday problem solving should focus on the development of problem interpretation (Berg and Calderone, 1994).

As it is interactive in nature, problem definition reflects those aspects of the self and context that are activated with respect to a specific problem unfolding at a specific moment in time. Problem definition is a complex, psychological, subjective reality, which, according to Berg et al. (1998), reflects the individual's goals and expectations (Bandura, 1986), determines the strategies to be used to meet these expectations and accomplish subjective goals (Vallacher and Wegner, 1987), affects the outcome attribution and meaning interpretation (Dodge, Pettit, McClaskey, and Brown, 1986), and induces the affective representation of the problem (Fleeson and Cantor, 1995).

A number of studies provide supportive evidence for the transactional approach to everyday problem solving. Sinnott (1989) showed that older adults' interpretation of Piagetian logical-combination problems, especially those experienced in real life (e.g., assigning relatives to sleeping locations), vary to a greater degree than do the interpretations of younger adults. Specifically, older adults tend to be more sensitive to social and interpersonal facets of the problem when compared with younger adults, who concentrate on the problem's logical aspects. Similarly, Laipple (1992) showed that older adults were less likely to interpret the situation of solving logical problems with the meaning intended by the experimenter; older adults tended to leave the logical confines of the problem and inject into the experimental situation more personal experience than did the younger adults. Chi and Ceci (1987) suggested that many types of problem solving appear to be directly influenced by the mental context the child brings to the task.

In their own work, Berg and colleagues (Berg and Calderone, 1994) registered a number of developmental characteristics of problem definition. First, they showed that, with age, there was a decrease in the frequency of task-oriented interpretations of problems and an increase in interpersonal, self, and mixed (e.g., task and self) interpretations. In their interpretation, researchers suggest that these findings correspond to the literature on the development of the self system, according to which changes of the self system involve movement away from a concrete and specific system to one that incorporates more abstract and interrelated psychological constructs (Harter, 1983). Second, Berg et al. (1998) studied the link between the problem definition and the selection of strategies for problem solving. In general, problem definition appears to be a more precise predictor of strategy use than does problem context. Specifically, individuals who defined a problem in terms of interpersonal concerns alone were more likely to report using strategies involving regulating or including others. On the contrary, individuals who defined a problem solely in terms of competence concerns were more likely to utilize strategies including independent action and less likely to engage others. Finally, the links between problem definition and strategy selection were not found to vary as a function of age.

Problem definition is very important to practical cognition. For example, a key difference between the results of Berg et al. (1998) and those of previous research is the importance that individuals placed on the social aspects of practical problem solving. Berg and colleagues found that the majority of individual problem definitions in any age group (pre-adolescents, college students, and adults) involved interpersonal concerns. These problem definitions, in turn, determined the selection of strategies that involved regulating or including others. Note that this interpretation differs significantly from the argument utilized in previous research. Earlier work typically assumed that reliance on others reflected ineffective problem solving because individuals exhibited dependence on others (e.g., Cornelius and Caspi, 1987; Denney and Palmer, 1981; Folkman et al., 1987). However, the reinterpretation of the role of social-dependent strategies suggests that using others to deal with everyday problems is a strategy rather well suited to particular problems (Baltes, 1997; Meacham and Emont, 1989).

2. Approaches to studying practical cognition

During the past two decades, there has been a growing interest (and in part a renewed interest) in nonacademic forms of cognition. Several distinct, but arguably overlapping, constructs have been proposed to capture this nonacademic form of cognition. One of these constructs is Sternberg's (1985a, 1997a) concept of practical cognition. Alternative related conceptualizations of nonacademic or practical cognition include *social cognition* (e.g., Cantor and Kihlstrom, 1987, Ford and Maher, 1998; Kihlstrom and Cantor, in press), *emotional cognition* (e.g., Goleman, 1995; Salovey and Mayer, 1990; Mayer, Salovey, and Caruso, in press), and *intrapersonal* and *interpersonal cognitions* (Gardner, 1983, 1993). Jones and Day (1997) noted the similarities among the various conceptualizations of nonacademic cognition. They suggested that practical, social, and emotional cognition share a focus on declarative and procedural knowledge, flexible knowledge-retrieval capabilities, and problem solving involving more than one correct interpretation or solution. We discuss the different conceptualizations of practical cognition and the methods researchers have used to study them.

2.1 Social cognition

Interest in the construct of social cognition has fluctuated since the concept was first introduced by Thorndike (1920). Thorndike defined social cognition as comprising the skills to understand others and to act or behave wisely in relation to others. He also distinguished social from abstract and mechanical forms of cognition. Several other definitions and expansions on Thorndike's definition followed. These expanded definitions included the skill to get along with others (Moss and Hunt, 1927), the skill to deal with people (T. Hunt, 1928), knowledge about people (Strang, 1930), ease with other people, insights into the states and traits of others (Vernon, 1933), and the skill to judge correctly the feelings, moods, and motivations of others (Wedeck, 1947). Wechsler's (1958) definition seemed to capture these various conceptualizations in the single definition of social cognition as one's facility in dealing with human beings.

Some researchers sought to understand the meaning of social cognition by studying people's implicit concepts or theories (e.g., Bruner, Shapiro, and Tagiuri, 1958; Cantor, 1978). In a study by Sternberg et al. (1981), discussed previously, experts and laypersons were asked to rate how characteristic various behaviors were of intelligent, academically intelligent, and everyday intelligent people. A factor of "social competence" emerged from the factor analyses of the ratings in each aspect of cognition.

More recently, Kosmitzki and John (1993) attempted to clarify some of the inconsistency in the literature regarding definitions of social cognition. They identified seven components that seemed to be most central to people's implicit conceptions of social cognition. The seven components included both cognitive elements (perspective taking, understanding people, knowing social rules, and openness to others) and behavioral elements (good at dealing with people, social adaptability, and interpersonal warmth). These implicit conceptions overlap, to some extent, with scientists' explicit theories, but suggest some additional aspects previously not included, such as interpersonal warmth and openness to others. Although these last two aspects have yet to be tested empirically, most studies have focused on some variation of the five remaining components (perspective taking, understanding people, knowing social rules, skill to deal with people, and social adaptability).

Throughout its history, the study of social cognition has periodically fallen out of favor with researchers. This lack of interest can be attributed to failed attempts to distinguish measures of social from measures of abstract cognition. The difficulty in

distinguishing social from academic or abstract cognition can be explained by efforts that focus primarily on cognitive aspects of social cognition and methods that rely heavily on verbal assessment. Researchers as early as Thorndike (1920) acknowledged the multidimensional nature of social cognition. Until recently, however, the approaches to studying social cognition have emphasized cognitive aspects, such as social perception (e.g., Chapin, 1942) and moral reasoning (e.g., Keating, 1978). In order to assess these cognitive dimensions, researchers relied, to a large extent, on verbal measures. Measures of behavioral aspects of social cognition also have relied somewhat on verbal forms of assessment (e.g., self-report). As becomes clear from a brief review of the literature, research efforts that consider behavioral and nonverbal measures of social cognition have had greater success in establishing discriminant validity from measures of abstract cognition than have the more cognitive, verbal measures of social cognition.

2.1.1 Cognitive-verbal measures of social cognition

Many approaches to understanding social cognition follow the tradition of cognition testing by developing instruments to assess individual differences in social cognition. One of the first and better known tests of social cognition was the George Washington Social Cognition Test (GWSIT; Moss, Hunt, Omwake, and Woodward, 1949). This test consists of a number of subtests that assess judgment in social situations, recognition of the mental states behind messages, memory for names and faces, observation of human behavior, and sense of humor. Early research with the GWSIT suggested that it could not be distinguished easily from abstract cognition (e.g., Thorndike and Stein, 1937).

A set of social-cognition tests emerged within the context of Guilford's (1967) Structure of Intellect Model of Cognition. Within Guilford's framework, social cognition is viewed as comprising those skills within the domain of behavioral operations. O'Sullivan, Guilford, and deMille (1965) developed tests to measure behavioral cognition, which they defined as the skill to judge people. More specifically, the tests measured the skill to decode social cues, including facial expressions, vocal inflections, posture, and gestures. In a study with 306 high-school students, O'Sullivan et al. (1965) found evidence that their factors of social cognition were distinct from measures of abstract cognitive skill. Later research, however, found contradictory results (e.g., Riggio, Messamer, and Throckmorton, 1991).

Riggio et al. (1991) administered several measures of social cognition and several measures of academic cognition to undergraduate students. Academic cognition was measured using the Shipley-Hartford Institute of Living Scale (Shipley, 1940), which measures verbal and abstract reasoning, and the vocabulary subscale of the WAIS-R (Wechsler, 1981). Measures of social cognition included four tests of the Factor Tests of Social Cognition (O'Sullivan and Guilford, 1976); Riggio's (1986, 1989) Social Skills Inventory (SSI), which assess six social communication skills (emotional expressivity, emotional sensitivity, emotional control, social expressivity, social sensitivity, and social control); and a social etiquette/tacit knowledge test that measured knowledge of appropriate behaviors in social situations. Riggio et al. found comparable intercorrelations within measures of both academic and social cognition as they did between measures of academic and social cognition. An exploratory factor analysis suggested two factors, one that included the Shipley-Hartford Abstract Reasoning scale and the Guilford measures and was labeled "abstract reasoning cognition," and the second that included the Shipley-Hartford Verbal scale and the SSI, which was labeled "verbal cognition." These findings suggested that academic and social cognition are overlapping domains. At the same time, these researchers found little evidence of convergent validity among the measures of social cognition, likely reflecting the

complexity of the construct and the various ways it has been operationalized in the literature.

Similar results were obtained by Keating (1978) using a different set of social-cognition measures. Keating administered the Social Insight Test (Chapin, 1967), which asks individuals to read about problem situations and to select the best from among four alternative interpretations of the situation; the Defining Issues Test (Rest, 1975), based on Kohlberg's (1963) theory of moral development; and the Social Maturity Index (Gough, 1966), which is a self-report measure of effective social functioning. Keating failed to find substantial intercorrelations among the social-cognition measures, and found no evidence, from either a multitrait-multimethod analysis or a factor analysis, that social cognition was distinct from academic cognition. All of Keating's measures, like those of Riggio et al. (1991), were verbal, which may have contributed to the inability to discriminate between abstract and social cognition.

2.1.2 Behavioral approaches to measuring social cognition

As a result of frustrations in trying to distinguish social from academic cognition, many researchers returned to Thorndike's (1920) definition and considered the behavioral as well as cognitive dimension of the construct. These researchers (e.g., Ford and Tisak, 1983; Frederickson, Carlson, and Ward, 1984) proposed that cognitive aspects of social cognition might expectedly be more closely associated with abstract cognition, whereas behavioral aspects would represent a more distinct construct.

A second set of tests to those of O'Sullivan et al. (1965) emerged from Guilford's (1967) Structure of Intellect model. These tests focused on behavioral rather than cognitive skills and defined social cognition as the skill to cope with people (Hendricks, Guilford, and Hoepfner, 1969). Hendricks et al. administered their tests to 252 high-schools students. Through principal-components analysis they identified factors that readily were interpretable as divergent production skills, and found that these factors were independent of behavioral cognition. These findings were later confirmed by Chen and Michael (1993).

A study by Ford and Tisak (1983) took the next step by distinguishing a behavioral measure of social cognition from academic cognition. The investigators conducted their study with more than 600 high-school students. Their measure of social cognition included self, peer, and teacher ratings of social competence, Hogan's (1969) empathy test, and a judgment of social competence from an individual interview. In addition, they obtained measures of verbal and math skill from school grades and standardized test scores. The measures of academic and social cognition were found to load on separate factors. They further found that the ratings of social competence and scores on the empathy scale were more predictive of interview ratings than were the measures of verbal and math skill. Ford and Tisak suggested that the difference between their findings and those of Keating (1978), reviewed earlier, were attributable to using a behavioral rather than a cognitive measure of social cognition.

A number of subsequent studies obtained findings consistent with Ford and Tisak (1983). Marlow (1986), for example, found that scores on several self-report measures of social cognition were unrelated to scores on measures of verbal and abstract cognition. Similarly, Frederickson et al. (1984) did not find significant correlations between ratings of interview behavior and measures of scholastic aptitude, achievement, or problem solving. However, Stricker and Rock (1990) did find a correlation between verbal skill and participants' skill to judge accurately a person and a situation from a videotaped interview.

Sticker and Rock (1990) administered a behavioral situational judgment test, the Interpersonal Competence Instrument (ICI), to 131 undergraduates, along with other measures of social cognition (e.g., peer and self-ratings, accuracy in decoding nonverbal communication) and measures of general skill (e.g., verbal comprehension, general reasoning). Using multidimensional scaling analysis, they found little evidence of convergent or discriminant validity among the measures of social cognition and general skill. Some of the social-cognition measures appeared to tap verbal skill, whereas others seemed to measure general reasoning skill. In contrast to the findings of Ford and Tisak (1983), these findings failed to support the hypothesis that behavioral measures of social cognition would be more distinguishable from measures of general academic cognition than would be verbal measures of social cognition.

Brown and Anthony's (1990) findings suggested that the constructs of social and academic cognition are distinct, but potentially interrelated. They evaluated the relationship of GPA and American College Test (ACT) English and Mathematics scores of college freshman to self and peer ratings of personality and behavioral dimensions of social cognition. Using a principal-components analysis, they identified three distinct components in their data, represented by an academic component, a peer-ratings component, and a self-ratings component. They concluded that social cognition could be separated from academic cognition. Social cognition as perceived by others was also distinct from one's own assessment of social skills. However, they also found that GPA was the best predictor of self and peer ratings on behavioral aspects of social cognition, suggesting a relationship between social skills and school performance.

2.1.3 Nonverbal approaches to measuring social cognition

In addition to behavioral approaches to measuring social cognition, researchers also sought to distinguish social from academic cognition by pursing nonverbal measures. Nonverbal approaches to measuring social cognition assess primarily nonverbal decoding skills (Archer, 1980; Archer and Akert, 1980; Barnes and Sternberg, 1989; Rosenthal, 1979; Rosenthal, Hall, DiMatteo, Rogers, and Archer, 1979; Sternberg and Smith, 1985). Rosenthal et al. developed the Profile of Nonverbal Sensitivity (PONS) test, which presents a single woman in a variety of poses. Participants are asked to decode the implicit signals being emitted, and to figure out which of two alternative descriptions better characterizes what the test taker has seen and/or heard. The PONS has been found to have weak to moderate correlations with other measures of social and cognitive competence (Halberstadt and Hall, 1980; Rosenthal et al., 1979).

Archer (1980; Archer and Akert, 1980) developed an alternative to the PONS test, called the Social Interpretation Test (SIT). The SIT presents participants with visual and auditory information regarding a social situation. For example, the participant might see a picture of a woman talking on the phone and hear a fragment of the woman's conversation. The participants are asked to judge whether the woman is talking to another woman or to a man. In another situation, participants are asked to judge whether a man and woman shown in a picture are strangers, acquaintances, or friends. Research using the SIT has focused primarily on the accuracy of participants' judgments based on verbal versus nonverbal information.

Using a task similar to the SIT, Sternberg and Smith (1985) developed a measure of decoding skills and assessed their relationship to other measures of social and cognitive cognition. They presented participants with two types of photographs. In one type, a man and woman were shown posing as if they were in a close relationship. Participants were asked to judge if the photograph depicted a real or a fake couple. In the second type, the picture showed a supervisor and his or her supervisee. Participants were asked to judge who of the two individuals was the supervisor. Accuracy was assessed as the

percentage of pictures the participant judged correctly. Participants were also given several measures of social and cognitive cognition including the PONS (Rosenthal et al., 1979); the Social Insight Test (Chapin, 1967); the George Washington Social Cognition Test (Moss et al., 1949); the Group Embedded Figures Test (Oltman, Raskin, and Witkin, 1971); and the Cattell Culture Fair Test of *g* (Cattell and Cattell, 1963). Nonverbal decoding accuracy only correlated significantly with performance on the Embedded Figures Test. Sternberg and Smith concluded that there was insufficient evidence to suggest that nonverbal decoding skills provided a valid measure of the construct of social cognition.

A subsequent study by Barnes and Sternberg (1989) was more successful. Participants were given the same set of pictures used in Sternberg and Smith (1985), one set portraying heterosexual couples and the other supervisors and supervisees. In addition to judging the pictures, participants were asked to rate their degree of confidence in their judgments; to indicate what features in each picture they used to make their judgement; to rate the importance of those features in their decision; and to assign a weight based on how much the feature was exhibited in the picture. Participants also completed several measures of social and academic cognition. They were assigned scores on *Social Competence* based on the 13 behaviors from the Sternberg et al. (1981) Social Competence Factor; *Situational Competence* based on the Social Competence Nomination Form (Ford, 1982); *Overall Social Competence* based on the Empathy Scale (Hogan, 1969) and the Self-Monitoring scale (Snyder, 1974); and *Overall Cognitive Cognition* based on educational background, school performance, and the Henmon-Nelson Test of Mental Skill (Nelson and Lamke, 1973). Barnes and Sternberg obtained significant correlations between accuracy at nonverbal decoding in the couples' task and all measures of social competence, except for situational competence. Decoding accuracy did not correlate with any of the cognitive cognition measures. There was, however, a correlation between the quantity of features identified by participants and cognitive cognition scores. These investigators concluded that the skill to accurately decode nonverbal communication is an indicator of social cognition.

Wong, Day, Maxwell, and Meara (1995) attributed previous failures to discriminate the two forms of cognition to the use of paper-and-pencil measures of social as well as academic cognition. Wong et al. conducted two studies to examine the relationships between cognitive and behavioral measures of social cognition and academic cognition. In the first study, they administered verbal, nonverbal, self-report and other-report measures of academic cognition, social perception (cognitive social cognition), and effectiveness in heterosexual interactions (behavioral social cognition) to undergraduate students. Using confirmatory factor analysis, they found that the model that best fit the data consisted of three separate factors: social perception, effectiveness in heterosexual interaction, and academic cognition. In the second study, they focused on three cognitive aspects of social cognition: social knowledge (knowledge of etiquette rules), social perception (the skill to understand the emotional states of others), and social insight (the skill to comprehend observed behaviors in a social context). The best-fitting model consisted of three factors: academic cognition, a combined social perception-social insight factor, and social knowledge. In their studies, Wong et al. were able to discriminate not only behavioral, but also cognitive aspects of social cognition from academic cognition.

Jones and Day (1997) attempted further to understand the cognitive and behavioral aspects of social cognition. They examined the relationship between two dimensions of social cognition, Crystallized Social Knowledge (declarative and procedural knowledge about familiar social events) and Social-Cognitive Flexibility (the skill to apply social knowledge to relatively novel problems). They proposed that these two dimensions of social cognition could be distinguished from academic problem solving, which depends

on fluid skills to solve novel, abstract problems that generally have a single, correct solution. They administered pictorial, verbal, self-report, and teacher-report measures of Crystallized Social Knowledge, Social-Cognitive Flexibility, and Academic Problem Solving to 169 high school students. In addition, they obtained a measure of social competence from the teachers. Confirmatory factor analyses of the correlation matrix among these measures indicated that the Social-Cognitive Flexibility factor could be discriminated from both Crystallized Social Knowledge and Academic Problem Solving, but that the latter were not discriminable from each other. They further found that all three factors were significantly related to social-competency ratings.

Although Jones and Day's (1997) findings suggest that there are different processes associated with solving novel social problems than those used to solve familiar social problems or novel academic problems, there are some limitations to their study. First, the sample (i.e., high school students) may represent individuals who are relative novices when it comes to social problem solving, such that their level of knowledge may reflect abstract concepts that are similar to academic-type problems. Individuals who have more expertise in social problems may have knowledge that is more distinct from academic problem-solving skill. Second, the method of measuring each of these factors may have contributed to the findings. Both Crystallized Social Knowledge and Academic Problem Solving involved items with one correct answer, whereas the measures of Social-Cognitive Flexibility asked respondents to provide their own interpretation, rate the importance of different social goals, and to identify the most effective solution to achieve the social goal. The similarity in the measurement format for the former two measures may have created an artificially higher validity estimate among them.

The limitations identified by Jones and Day (1997) are some of the concerns that Kihlstrom and Cantor (in press) raise about relying on psychometric approaches to study social cognition. Cantor, Kihlstrom and colleagues (Cantor and Harlow, 1994; Cantor and Kihlstrom, 1987; Kihlstrom and Cantor, in press) take a social-cognition view of personality. These researchers do agree that social behavior is intelligent because it is mediated by cognitive processes such as perception, memory, reasoning, and problem solving. They argue that psychometric approaches to understanding social cognition inappropriately focus on how much social cognition a person has rather than what social cognition the person possesses. Individual differences in social behavior can be attributed to differences in knowledge and strategies needed to accomplish social tasks.

Cantor and Harlow (1994) proposed that intelligent behavior involves attunement to the consequences of one's actions, the implications of those consequences for other goals, and the goal-fulfilling potentials of different situations. Attunement allows for flexibility in terms of what tasks to pursue, where and when opportunities are present to work on various tasks, and how to pursue the tasks. Therefore, attunement and flexibility are critical aspects of personality and cognition, allowing individuals successfully to pursue goal and solve problems. Cantor and Harlow argued that, due to the varied tasks and settings in which individuals behave, it is difficult to obtain a general, stable measure of social cognition.

Rather than developing instruments to assess individual differences in social cognition, Cantor and her colleagues (see Kihlstrom and Cantor, in press) have chosen to study the cognitive processes that support intelligent social behavior. They focus on *life tasks* as their unit of analysis for studying social cognition. Life tasks are identified by the individual as meaningful and serve to organize one's daily activities. They allow researchers to observe the skill of people to solve problems of a social nature and the knowledge they have of how to solve them. Life tasks include things like making friends, finding a spouse, establishing a career, and getting good grades. Cantor and her colleagues have chosen to focus on periods of transition (e.g., from high school to college) to

observe individual differences in life tasks (see Kihlstrom and Cantor, in press). They have found that people formulate action plans, monitor their progress, and assess the outcomes of their actions. They draw on their autobiographical memory to evaluate various causes of those outcomes and alternative courses of action. When their pursuit of a life task is obstructed, they are able to alter plans or choose new plan of action. As we discuss in a later part of this section, the processes identified by Cantor and her colleagues are consistent with the metacomponents identified by Sternberg (1985a) as underlying successful cognition.

Unfortunately, recent efforts to define and measure social cognition have not led to any substantial improvement in our understanding of the construct. There appear to be as many definitions and operationalizations of social cognition as there are researchers. The definitions of social cognition reference dimensions such as social perception, social knowledge, social insight, empathy, social memory, and social adaptation. Furthermore, there is little consistency regarding the relationships among measures of social cognition or their relations to measures of academic cognition. Although we acknowledge Cantor and Harlow's (1994) concern regarding the difficulty in measuring social cognition, the construct of tacit knowledge, elaborated upon in other chapters, represents an attempt to quantify context-specific knowledge that is an aspect of practical cognition. Tests of tacit knowledge have been successful in predicting performance (behavioral outcomes) and discriminating practical from abstract or academic cognition (e.g., Sternberg et al., 1993; Sternberg et al., 1995). Before considering the measurement of practical cognition, we discuss another related construct, that of emotional cognition.

2.2 Emotional cognition

Research and theorizing on the construct of emotional cognition (also called, perhaps inappropriately, "emotional intelligence") has a much shorter history in comparison to social cognition. According to Mayer, Salovey, and Caruso (in press), the history of emotional cognition research spans less than a decade. As such, the number of definitions of and approaches to studying emotional cognition are delineated more readily.

Mayer et al. (in press) distinguished between two general models of emotional cognition. *Skill models* view emotional cognition as the intersection of cognition and emotion. *Mixed models* define emotional cognition as a combination of mental skill and personality traits. We talk first about the mixed models (e.g., Bar-On, 1997; Goleman, 1995) and their associated measures of emotional cognition. Then we discuss the work of Mayer and Salovey as they attempt to characterize emotional cognition as distinct from personality.

Goleman (1995) brought popular attention to the concept of emotional cognition. He argued, as other researchers have (e.g., Gardner, 1983; Sternberg, 1997a), that cognitive-ability tests and similar tests (e.g., SATs) fail to predict accurately who will succeed in life. Goleman suggested that part of the 80% variance in success unaccounted for by overall cognitive ability could be explained by other characteristics, one of which is emotional cognition. He defined emotional cognition as including "skills such as being able to motivate oneself and persist in the face of frustrations; to control impulses and delay gratification; to regulate one's moods and keep distress from swamping the skill to think; to empathize and to hope" (p. 34). Although Goleman did not point to any specific test of emotional cognition, he cited support for the construct in research on related factors, such as empathy and ego resilience, which suggests that emotional cognition is distinct from overall cognitive ability.

Interestingly, Mayer et al. (in press) cited a study by Davies, Stankov, and Roberts (1998) that used a scale Goleman created to measure emotional cognition. The items consisted of hypothetical situations to which individuals responded. Davies et al. (as reported in Mayer et al.) found that Goleman's measure correlated with self-reported empathy and emotional control. Mayer et al. noted that it is not clear whether Goleman's scale was intended for empirical use, so the findings of Davies et al. are tentative.

A more measurement-based approach is represented by Bar-On (1997), who defined emotional cognition as all non-cognitive skills and competencies that enable one to cope successfully with life. Bar-On identified five broad areas of skills or competencies, and within each, more specific skills that appear to contribute to success. These include intrapersonal skills (emotional self-awareness, assertiveness, self-regard, self-actualization, independence); interpersonal skills (interpersonal relationships, social responsibility, empathy); adaptability (problem solving, reality testing, flexibility); stress management (stress tolerance, impulse, control); and general mood (happiness, optimism). According to Mayer et al. (in press), Bar-On's model combines skills that can be characterized as mental skills (e.g., problem solving) and others that can be considered personality traits (e.g., optimism), thus making it a mixed model.

Bar-On (1997) developed the Emotional Quotient Inventory (EQ_i) based on his broad-based model of non-cognitive skills. Thirteen subscales of the EQ_i were identified, roughly corresponding to the specific skills in his model. These subscales were found to be highly intercorrelated, and thus a single test score is computed. Bar-On has found that scores on his test correlate negatively with measures of negative affect (Beck Depression Inventory; Beck, Ward, Mendelson, Mock, and Erbaugh, 1961; Zung Self-Rating Depression Scale); positively with measures of positive affect (e.g., emotional stability; extraversion); and nonsignificantly with measures of general cognition (e.g., WAIS-R; Wechsler, 1981). Again, it is clear from these results that Bar-On's EQ_i measures aspects of personality, and possibly mental skill. Because the measure is one of self-report, it is difficult to assess how generalizable the results would be to behavior.

Initial theorizing by Salovey and Mayer (1990) also related emotional cognition to personality factors such as warmth and outgoingness. But in the time since, they have argued that these personality factors are distinct from emotional cognition. They consider the latter to be more strictly a skill (Mayer and Salovey, 1997; Mayer et al., in press). They define emotional cognition as the skill to recognize the meanings of emotions and to use that knowledge to reason and solve problems. They have proposed a framework of emotional cognition to organize the various skills involved in the adaptive processing of emotionally relevant information.

Emotional cognition consists of four main classes of skills. These skills pertain to (1) the accurate appraisal and expression of emotions in oneself and in others, (2) assimilation of emotional experience into cognition, (3) recognition, understanding, and reasoning about emotions, and (4) the adaptive regulation of emotions in oneself and in others (Mayer et al., in press; Salovey and Mayer, 1994).

Mayer and Salovey (1993) offered several mechanisms underlying emotional cognition that suggest its association with mental skills. First, emotions are associated with thought processes—certain emotions may increase thoughts and direct attention to certain tasks. Second, the effective regulation of emotions may be related to other skills, such as empathy and openness. Third, research on *alexithymia* (the inability to appraise and verbally express emotions) suggests possible disconnections between areas of the brain that prohibit the integration of thoughts and emotions.

Mayer and Salovey (1997; Mayer, Caruso, and Salovey, in press) have developed their own test of emotional cognition, called the Mutifactor Emotional Cognition Scale (MEIS). It consists of twelve skill measures that fall into the four classes of skills identified

above (perception, assimilation, understanding, and managing emotions). Perception is measured by presenting various stimuli, including faces, abstract designs, music, and stories, and asking people to judge the emotional content reflected in those stimuli. Assimilation is measured by *Synesthesia Judgments* (describing emotional sensations and their relations to other sense modalities) and *Feeling Biases* (judgment of how the individual feels toward a fictional person). Understanding is measured by *Blends* (the skill to blend emotions; e.g., Optimism most closely combines which two emotions?); *Progressions* (understanding how emotional reactions progress over time); *Transitions* (understanding how emotions flow from one to another); and *Relativity* (estimating the feelings of people depicted in a conflictual social encounter). Finally, managing emotions is measured in reference to others and to oneself. Managing feelings of others is measured using brief vignettes about fictional people in need of assistance and asking the respondent to rate the effectiveness of alternative courses of action. Managing feelings of self is measured similarly, but the vignettes describe emotional problems that the individual might encounter.

Mayer et al. (1998) validated the MEIS with 503 adults and 229 adolescents. From a factor analysis of the MEIS, Mayer et al. identified three primary factors corresponding to *Perception*, *Understanding*, and *Managing* emotion, and a higher order, general factor of *Emotional Cognition* (g_{ei}; Mayer et al., in press). General emotional cognition correlated significantly with a measure of verbal cognition (the Army Alpha vocabulary scale; Yerkes, 1921) and a measure of self-reported empathy (Caruso and Mayer, 1997). The investigators also found that the emotional cognition of adults was higher than that of adolescents, suggesting age-related changes. Of the three specific factors, *Understanding* correlated most highly with verbal cognition, followed by *Managing* emotions and then *Perception*. These investigators concluded that emotional cognition can be characterized appropriately as a mental skill because their results follow the patterns of other well-established measures of cognition. The specific skills in the MEIS are intercorrelated, scores on the MEIS develop with age as do scores on other standard cognition tests, and emotional cognition overlaps, to some extent, with traditional cognition.

Schutte et al. (1998) developed their own measure of emotional cognition based on Salovey and Mayer's (1990) model. Their 33-item self-report measure correlated significantly with eight theoretically related constructs, including awareness of emotion, outlook on life, depressed mood, skill to regulate emotions, and impulsivity. They also showed differences on their measure with groups expected to differ in emotional cognition (e.g., psychotherapists and prisoners, men and women). They further showed that scores on indices of emotional cognition were predictive of end-of-year grade-point averages of college freshman, but were unrelated to SAT or ACT scores. Finally, they found that of the big five personality traits, emotional cognition related significantly only to openness to experience.

There appears to be some support for both the construct of social cognition and that of emotional cognition. As yet, there have been no direct efforts aimed at distinguishing social from emotional cognition, and often the two are treated interchangeably. However, there is evidence to suggest that both social and emotional cognition overlap, to some extent, with abstract, academic cognition. This interdependence is not surprising if we take the position that similar mental processes are employed in solving problems of a social, emotional, or academic nature. Sternberg's (1997a) theory of successful cognition, and the triarchic theory subsumed within it, specifies these processes and their relation to successful performance of everyday tasks. Before considering his theory, we briefly review some alternative frameworks of competence or cognition that provide a different perspective on social, emotional, and even abstract cognition.

2.3 Comprehensive frameworks of skills

Some researchers have attempted to define nonacademic forms of cognition within broader models of personal competence (Greenspan, 1981; Greenspan and Driscoll, 1997; Greenspan and Granfield, 1992) or human functioning (D. Ford, 1987, 1994; M. Ford and D. Ford, 1987; M. Ford and Maher, 1998). We briefly review two of these frameworks here.

2.3.1 Greenspan and driscoll's model of personal competence

Greenspan and his colleagues (Greenspan, 1981; Greenspan and Driscoll, 1997; Greenspan and Granfield, 1992) view personal competence as comprising the skills involved in attaining goals and solving problems, whereas cognition refers to the subcomponent of these skills involved in thinking and understanding. A recent version of their model (Greenspan and Driscoll, 1997) consists of four broad domains of competence: physical competence, affective competence, everyday competence, and academic competence. These broad domains are further divided into eight subdomains. Physical competence consists of organ (e.g., vision, heart functioning) and motor competence (e.g., strength, coordination). Affective competence consists of temperament (e.g., emotionality, distractibility) and character (e.g., gregariousness, social orientation). Everyday competence includes practical cognition (i.e., the skill to think about and understand problems in everyday settings) and social cognition (i.e., the skill to think about and understand social problems). Academic competence involves conceptual cognition (i.e., the skill to think about and understand problems of an academic or abstract nature) and language (i.e., the skill to understand and participate in communications).

Greenspan and Driscoll's (1997) model takes into account Cantor and Kihlstrom's (1989) suggestion that social cognition forms a link between cognition and personality. The tendency to view personality as a disposition and cognition as a skill has led most researchers to treat the constructs as separate. The Greenspan-Driscoll model recognizes that social competence consists of both intellective and nonintellective components.

2.3.2 The living systems framework

In the Living Systems Framework (LSF) of human functioning and development (D. Ford, 1987, 1994; M. Ford and D. Ford, 1987), cognition is viewed as the effective pursuit of goals within some setting or domain of activity (M. Ford and Maher, 1998). The key aspect of the LSF is the *behavior episode*, a context-specific, goal-directed pattern of behavior. Everyday life consists of a continuous series of behavior episodes. Behavior episodes can involve motor or communicative activity, information seeking, or thought processes. Multiple behavior episodes form a *behavior episode schema* (BES) that directs attention and guides thoughts, feelings, and actions, and consists of both declarative and procedural knowledge. The combination of a number of BESs allows for flexibility in dealing with various types of everyday problems, which is considered a major component of social and practical cognition (M. Ford, 1986).

Neither the Greenspan model nor that of M. Ford and Maher (1998) seems to capture emotional cognition as defined by Mayer et al. (in press). It is likely that Greenspan and Driscoll would consider emotional cognition as they view social cognition, that is, at the intersection of personality and cognition. Both these models, and the approaches to social and emotional cognition discussed above, recognize the importance of nonacademic or nontraditional cognition in determining success in life. This view also forms the basis of Sternberg's (1997a) aptly named theory of successful cognition.

2.4 Sternberg's theory of successful cognition

Consistent with Greenspan and Driscoll's distinction between academic and everyday competence is Sternberg's (1985a) distinction between academic and practical cognition. Practical cognition, however, is part of a more comprehensive theory of successful cognition (Sternberg, 1997a). According to the theory, successful cognition is the skill to achieve success in life, given one's personal standards, within one's sociocultural context. One's skill to achieve success depends on one's capitalizing on one's strengths and correcting or compensating for one's weaknesses through a balance of analytical, creative, and practical skills in order to adapt to, shape, and select environments.

The theory of successful cognition, first introduced in second section, serves as the basis for the work described throughout this book on practical cognition and tacit knowledge. We describe in greater detail in this section the main components of the theory. Then we describe a measure designed to assess these components, including the skill to apply knowledge to real-world, practical problems.

Sternberg's theory of successful cognition (Sternberg, 1988, 1997a) seeks to explain in an integrative way the relationship between cognition and (1) the internal world of the individual, or the mental mechanisms that underlie intelligent behavior; (2) experience, or the mediating role of one's passage through life between the internal and external worlds; and (3) the external world of the individual, or the use of cognitive mechanisms in everyday life in order to attain a functional fit to the environment. These three parts of the theory are referred to respectively as the componential subtheory, the experiential subtheory, and the contextual subtheory.

The componential subtheory. The componential subtheory seeks to elucidate the mental processes that underlie intelligent behavior by identifying three basic kinds of information-processing components, referred to as metacomponents, performance components, and knowledge-acquisition components.

Metacomponents are higher order, executive processes used to plan what one is going to do, to monitor it while one is doing it, and evaluate it after it is done. These metacomponents include (1) recognizing the existence of a problem, (2) deciding on the nature of the problem confronting one, (3) selecting a set of lower order processes to solve the problem, (4) selecting a strategy into which to combine these components, (5) selecting mental representation on which the components and strategy can act, (6) allocating one's mental resources, (7) monitoring one's problem solving as it is happening, and (8) evaluating one's problem solving after it is done.

Performance components are lower order processes that execute the instructions of the metacomponents. These components solve the problems according to the plans laid out by the metacomponents. Whereas the number of metacomponents used in the performance of various tasks is relatively limited, the number of performance components is probably quite large, and many are relatively specific to a narrow range of tasks (Sternberg, 1985a). Inductive reasoning tasks such as matrices, analogies, series completion, and classifications involve a set of performance components that provide potential insight into the nature of the general factor of cognition. That is, induction problems of these kinds show the highest loading on the general cognition factor, or *g* (Jensen, 1980; Snow and Lohman, 1984; Sternberg and Gardner, 1982). The main performance components of inductive reasoning are encoding, inference, mapping, application, comparison, justification, and response.

Knowledge-acquisition components are used to learn how to do what the metacomponents and performance components eventually do. Three knowledge-acquisition components seem to be central in cognitive functioning: (1) selective encoding, (2) selective combination, and (3) selective comparison.

Selective encoding involves sifting out relevant information from irrelevant information. When new information is presented in natural contexts, relevant information for one's given purpose is embedded in the midst of large amounts of purpose-irrelevant information. A critical task for the learner is to sift the "wheat from the chaff," recognizing just what among the pieces of information is relevant for one's purposes (see Schank, 1990).

Selective combination involves combining selectively encoded information in such a way as to form an integrated, plausible whole. Simply sifting out relevant from irrelevant information is not enough to generate a new knowledge structure. One must know how to combine the pieces of information into an internally connected whole (see Mayer and Greeno, 1972).

Selective comparison involves relating new information to old information already stored in memory. It is not enough to encode and combine new information; the information has to be tied to some preexisting knowledge base. A good selective comparer recognizes how existing knowledge can be brought to bear on the present situation. A poor selective comparer does not readily see the relations between existing and new information. For example, a competent lawyer looks for past precedents, a competent doctor for old cases that shed light on new ones.

The various components of cognition work together. Metacomponents activate performance and knowledge-acquisition components. These latter kinds of components in turn provide feedback to the metacomponents. Although one can isolate various kinds of information-processing components from task performance using experimental means, in practice, the components function together in highly interactive ways, and are not readily isolated. Thus, diagnosis as well as instructional interventions needs to consider all three types of components in interaction rather than any one kind of component in isolation. But understanding the nature of the components of cognition is not, in itself, sufficient to understand the nature of cognition because there is more to cognition than a set of information-processing components. One could scarcely understand all of what it is that makes one person more intelligent than another by understanding the components of processing on, say, a cognition test. The other aspects of the triarchic theory address some of the other aspects of cognition that contribute to individual differences in observed performance, outside testing situations as well as within them.

The experiential subtheory. Components of information processing always are applied to tasks and situations with which one has some level of prior experience (including the null level). Hence, these internal mechanisms are closely tied to one's experience. According to the experiential subtheory, the components are not equally good measures of cognition at all levels of experience. Assessing cognition requires one to consider not only components but also the level of experience at which they are applied.

According the experiential subtheory, cognition is best measured at those regions of the experiential continuum that involve tasks or situations that are either relatively novel, on the one hand, or in the process of becoming automatized, on the other.

Several sources of evidence converge on the notion that skill to deal with relative novelty is a good way of measuring cognition. Davidson and Sternberg (1984) found that gifted children had greater insight to deal with novel problems than did nongifted children. Research on fluid cognition, which is a kind of cognition involved in dealing with novelty (see Cattell, 1971), suggests that tests that measure the skill to deal with novelty fall relatively close to the so-called general factor of cognition (Snow and Lohman, 1984).

There are also converging lines of evidence that automatization skill is a key aspect of cognition. Sternberg (1977) found that the correlation between performance on an analogy problem and measure of general cognition increased with practice. The first stage of Ackerman's (1987; Kanfer and Ackerman, 1989) model of automatization also is related to cognition. Theorists such as Jensen (1982) and Hunt (1978) attribute the correlation between such tasks as choice reaction time and letter matching to the relation between speed of information processing and cognition. An alternative explanation is that some of the correlation is due to the effects of automatization of processing.

The skill to deal with novelty and the skill to automatize information processing are interrelated. If one is able to automatize well, one has more resources left over for dealing with novelty. Similarly, if one is well able to deal with novelty, one has more resources left over for automatization.

The contextual subtheory. According to the contextual subtheory, intelligent thought is directed toward one or more of three behavioral goals: (a) adaptation to an environment, (b) shaping of an environment, or (c) selection of an environment. These three goals may be viewed as the functions toward which cognition is directed. Cognition is not aimless or random mental activity that happens to involve certain components of information processing at certain levels of experience. Rather, it is purposefully directed toward the pursuit of these three global goals, all of which have more specific and concrete instantiations in people's lives.

Most intelligent thought is directed toward the attempt to adapt to one's environment. The requirements for adaptation can differ radically from one environment to another—whether environments are defined in terms of families, jobs, subcultures, or cultures. According to the triarchic theory, and the contextual subtheory in particular, the processes, experiential facets, and functions of cognition remain essentially the same across contexts, but the particular instantiations of these processes, facets, and functions can differ radically. Thus, the content of intelligent thought and its manifestations in behavior will bear no necessary resemblance across contexts. To understand cognition, one must understand it, not only in relation to its internal manifestations in terms of mental processes and its experiential manifestations in terms of facets of the experiential continuum, but also in terms of how thought is intelligently translated into action in a variety of different contextual settings. The difference in what is considered adaptive and intelligent can extend even to different occupations within a given cultural milieu.

Shaping of the environment is often used as a backup strategy when adaptation fails. If one is unable to change oneself to fit the environment, one may attempt to change the environment to fit oneself. Shaping, however, is not always used in lieu of adaptation—it may be used before adaptation is tried. In science, the distinction can be made between those who set the paradigms (shape) and those who follow them (adapt) (see Sternberg, 1999).

Selection involves renunciation of one environment in favor of another. Selection is sometimes used when both adaptation and shaping fail. Failure to adjust to the demands of a work environment, or to change the demands to fit one's interest, values, expectations, or skills, may result in a decision to seek a new job. But selection is not always used as a last resort. It may reflect an intelligent person's recognition that a situation is not suitable and that no attempt to change oneself would improve the fit.

Adaptation, shaping, and selection are functions of intelligent thought as it operates in context. It is through adaptation, shaping, and selection that the components of cognition, as employed at various levels of experience, become actualized in the real world. This is the definition of practical cognition used by Sternberg and his colleagues (e.g., Sternberg, 1997a; Sternberg and Wagner, 1986).

2.4.1 Sternberg triarchic abilities test

A measure was developed to assess the components of Sternberg's theory (Sternberg, 1985a, 1988). The Sternberg Triarchic Abilities Test (STAT; Sternberg, 1991a, 1991b, 1993) measures three domains of mental processing (analytical, creative, and practical), which reflect the subtheories outlined above. Analytical questions address the skill to learn from context and reason inductively (i.e., the relation of cognition to the internal world). Creative questions address the skill to cope with novelty (i.e., the relation of cognition to experience). And practical questions address the skill to solve real-world, everyday problems (i.e., the relation of cognition to the external world).

The current version of the STAT (1993) has nine four-option multiple-choice subtests, each consisting of four items, plus three essays. The nine multiple-choice subtests represent a crossing of three kinds of process domains (analytical, creative, and practical) with three major content domains (verbal, quantitative, and figural). The three essays assess performance in analytical, creative, and practical domains. We describe each of the subtests below, organized around the process domains.

There are four analytical subtests of the STAT, one for each content area (multiple-choice verbal, multiple-choice quantitative, multiple-choice figural, and essay). Traditional verbal skill tests (e.g., synonym/antonym tests) correlate highly with overall cognitive ability (see Sternberg and Powell, 1983), but they are more measures of achievement than of skill. In other words, they emphasize the products over the process of learning. Analytical-verbal skills are measured in the STAT by assessing the skill to learn from context. Vocabulary is viewed as a proxy for the skill to pick up information from relevant context (see Sternberg, 1987). The analytical-quantitative section consists of items that measure inductive reasoning skill in the numerical domain. The analytical-figural items similarly measure inductive reasoning skill with either figure classification or figure analogy problems. In the figure classification test, the examinee must indicate which figure does not belong with the others. The four analytical subtests are described below:

1. Analytical-Verbal (neologisms). Students see a novel word embedded in a paragraph, and have to infer its meaning from the context.
2. Analytical-Quantitative (number series). Students have to say what number should come next in a series of numbers.
3. Analytical-Figural (matrices). Students see a figural matrix with the lower right entry missing, and have to say which of the options fits into the missing space.
4. Analytical-Essay. Students are required to analyze the advantages and disadvantages of having police or security guards in a school building.

The creative portion of the STAT also consists of four subtests (multiple-choice verbal, multiple-choice quantitative, multiple-choice figural, and essay). The creative-verbal questions require counterfactual reasoning and attempt to assess the skill to think in relatively novel ways. In the creative-quantitative questions, symbols are used in place of certain numbers requiring the examinee to make a substitution. The creative-figural items require the examinee to complete a series in a domain separate from the one in which they inferred the completion rule. The four creative subtests are described below:

5. Creative-Verbal (novel analogies). Students are presented with verbal analogies preceded by counterfactual premises (e.g., money falls off trees), and must solve the analogies as though the counterfactual premises were true.

6. Creative-Quantitative (novel number operations). Students are presented with rules for novel number operation (e.g., flix, for which numerical manipulations differ depending upon whether the first of two operands is greater than, equal to, or less than the second). Students have to use the novel number operations to solve presented math problems.
7. Creative-Figural (novel series completion). Students are first presented with a figural series that involves one or more transformations; they then must apply the rule of the original series to a new figure with a different appearance, to complete a new series.
8. Creative-Essay. Students are required to describe how they would reform their school system to produce an ideal one.

Finally, the practical portion of the STAT is designed to assess the skill to apply knowledge to problems with practical relevance. Practical-verbal items require the examinee to answer everyday inferential reasoning problems. Practical-quantitative items require the examinee to reason quantitatively with practical everyday problems of the kind he or she might face in everyday life. Items in the practical-figural portion require the skill to plan a route efficiently, given the information in a map or diagram. The four practical subtests are described below:

9. Practical-Verbal (everyday reasoning). Students have to solve a set of everyday problems in the life of an adolescent (e.g., what to do about a friend who seems to have a substance-abuse problem).
10. Practical-Quantitative (everyday math). Students have to solve math problems based on scenarios requiring the use of math in everyday life (e.g., buying tickets for a ballgame or making chocolate chip cookies).
11. Practical-Figural (route planning). Students are presented with a map of an area (e.g., an entertainment park), and have to answer questions about navigating effectively through the area depicted by the map.
12. Practical-Essay. Students are required to specify a problem in their life, and to state three practical solutions for solving it.

The multiple-choice questions are scored using an answer key. The essays are scored by trained raters according to the extent to which the answer reflects analytical, creative, and practical thinking. In a pilot use of the STAT (Sternberg and Clinkenbeard, 1995), a variety of skill tests were administered to 64 participants. The other tests used were the Terman Concept Mastery Test (primarily a test of crystallized skills), the Watson-Glaser Critical Thinking Appraisal (a verbal test of critical thinking), the Cattell Culture Fair Test of *g* (primarily a test of fluid skills), and a homemade test of insight problems (adapted from Sternberg, 1986). Respective correlations of the STAT with these tests were, for the analytical .49, .50, .50, and .47 (all significant); for the creative, .43, .53, .55, and .59 (all significant); and for the practical .21, .32, .36, and .21 (the second and third significant). Of the three processing domains measured by the STAT, the one that correlated the least with more traditional measures of general cognition was practical skill.

In a subsequent study (Sternberg, Ferrari, Clinkenbeard, and Grigorenko, 1996; Sternberg, Grigorenko, Ferrari, and Clinkenbeard, 1999), the STAT was administered to 324 children around the United States and in some other countries who were identified by their schools as gifted by any standard whatsoever. Children were selected for a summer psychology program at Yale (college-level) if they fell into one of five skill groupings: high analytical, high creative, high practical, high balanced (high in all three skills), or low balanced (low in all three skills). Students who came to Yale were then divided into four instructional groups. Students in all four instructional groups used the same introductory-psychology textbook (a preliminary version of Sternberg [1995b])

and listened to the same psychology lectures. What differed among them was the type of afternoon discussion section to which they were assigned. They were assigned to an instructional condition that emphasized either memory, analytical, creative, or practical instruction. For example, in the memory condition, they might be asked to describe the main tenets of a major theory of depression. In the analytical condition, they might be asked to compare and contrast two theories of depression. In the creative condition, they might be asked to formulate their own theory of depression. In the practical condition, they might be asked how they could use what they had learned about depression to help a friend who was depressed.

Students in all four instructional conditions were evaluated in terms of their performance on homework, a midterm exam, a final exam, and an independent project. Each type of work was evaluated for memory, analytical, creative, and practical quality. Thus, all students were evaluated in exactly the same way.

Sternberg et al. (1996) performed a principal-components factor analysis and found a weak general factor, suggesting that the general factor of cognition is probably relevant only when a fairly narrow range of skills is measured, as is typically the case with conventional tests. They found that testing format had a large effect on results: multiple-choice tests tended to correlate with other multiple-choice tests, almost without regard to what they measure. Essay tests showed only weak correlations with multiple choice, however. These investigators further found that after they controlled for modality of testing (multiple-choice versus essay), the correlations between the analytical, creative, and practical sections were very weak and generally nonsignificant, supporting the relative independence of the various skills. All three skill tests—analytical, creative, and practical—significantly predicted course performance. When multiple-regression analysis was used, at least two of these skill measures contributed significantly to the prediction of each of the measures of achievement. Perhaps as a reflection of the difficulty of deemphasizing the analytical way of teaching, one of the significant predictors was always the analytical score. Most importantly, there was an aptitude-treatment interaction whereby students who were placed in instructional conditions that better matched their pattern of skills outperformed students who were mismatched. In other words, when students are taught in a way that fits how they think, they do better in school. Children with creative and practical skills, who are almost never taught or assessed in a way that matches their pattern of skills, may be at a disadvantage in course after course, year after year.

Thus the results of the studies involving the STAT suggest that the theory of successful cognition is valid not just in its parts but as a whole. Moreover, the results suggest that the theory can make a difference not only in laboratory tests, but in school classrooms as well.

More recently, the triarchic theory of cognition was tested with an adult population, using alternative measures of analytic, creative, and practical cognition from the STAT questions described above. Grigorenko and Sternberg (in press) administered measures of analytical, creative, and practical cognition to 452 women and 293 men between the ages of 26 and 60 in a large industrial city in Russia. The environment in Russia is characterized by financial, institutional, political, and societal uncertainty and instability. The investigators hypothesized that, in such environments, practical and creative cognition would play as important a role, if not a greater role, in the successful adaptation to the changing social context.

Grigorenko and Sternberg measured analytical skill using the *Series* and the *Matrices* subtests of the Test of g: Culture Fair, Level II (Cattell, 1940; Cattell and Cattell, 1973) to measure fluid skills. A test of crystallized cognition was adapted from existing traditional tests of analogies and synonyms/antonyms used in Russia. Creative cognition

was measured by asking participants to describe the world through the eyes of insects and to describe who might live and what might happen on a fictitious planet called "Priumliava." Responses were rated for novelty, quality, and sophistication. For practical cognition, participants were asked to report their practical skills in the social domain (e.g., effective and successful communication with other people), in the family domain (e.g., how to fix household items, how to run the family budget), and in the domain of effective resolution of sudden problems (e.g., organizing something that has become chaotic). The participants were also asked to respond to 4 vignettes, based on themes of (1) how to maintain the value of one's savings; (2) what to do when one makes a purchase and discovers that the item one has purchased is broken; (3) how to locate medical assistance in a time of need; and (4) how to manage a salary bonus one has received for outstanding work. Participants were asked to select the best option among five presented for each vignette. The most frequently chosen option was used as the keyed answer. Finally, self-report measures of physical and mental health were used to assess successful adaptation. Participants received a summary score on their physical health based on reports of chronic illness and other debilitation injuries or diseases. They also completed the Beck Anxiety Scale (BAS, Beck, Epstein, Brown, and Steer, 1988) and the Beck Depression Inventory (BDI, Beck et al., 1961), as well as five items that measured their self-efficacy for adaptation.

Grigorenko and Sternberg found that practical cognition consistently predicted self-reported adaptive functioning on all indicators, with higher practical cognition associated with better physical and mental health. Analytical cognition was associated with lower anxiety and higher self-efficacy on two items. Creative cognition was marginally associated with poorer physical health, but lower anxiety. When the data was analyzed separately by gender, creative skill was found to associate with lower anxiety for women, but poorer physical health and lower self-efficacy for men. The results suggest that both analytical and practical cognition have a positive effect on adaptive functioning. There is inconclusive evidence regarding the role of creative cognition.

2.4.2 Measures of practical cognition

In addition to the STAT and self-report questions, Sternberg and his colleagues have developed measures targeted specifically at practical cognition (see Sternberg et al., 1993; Sternberg et al., 1995; Wagner, 1985). Practical cognition is viewed as relevant to successful performance of everyday problems, whether the problems are of a social, emotional, or task-related nature. Therefore, measures of practical cognition hold promise for elucidating some of the unexplained portion of success that have not been accounted for by traditional cognition tests. Sternberg and his colleagues have taken a knowledge-based approach to measuring practical cognition. *Tacit knowledge*, as an aspect of practical cognition, is experience-based knowledge relevant to solving practical problems. As such, tacit knowledge can pertain to social or emotional information. Therefore, tacit knowledge may provide a common approach to understanding various forms of nonacademic cognition. In the sections that follow, we delineate further the construct of tacit knowledge, describe methods of measuring tacit knowledge, and review a program of research that provides growing support for the validity of tacit knowledge, and subsequently, practical cognition.

3. Understanding practical cognition: The role of tacit knowledge

What distinguishes people who are more successful from those who are less successful in their everyday lives? Sternberg and his colleagues (Sternberg et al., 1993; Sternberg et al., 1995; Wagner and Sternberg, 1985; Wagner, 1987) have taken a knowledge-based approach to addressing this question. They have found in their research that much of the knowledge needed to succeed in real-world tasks is tacit. It is acquired while performing everyday activities, but typically without conscious awareness of what is being learned. And although people's actions may reflect their knowledge, they may find it difficult to articulate what they know. The notion that people acquire knowledge without awareness of what is being learned is reflected in the common language of the workplace as people speak of "learning by doing" and of "learning by osmosis." Terms like *professional intuition* and *professional instinct* further imply that the knowledge associated with successful performance has a tacit quality.

The term *tacit knowledge*, introduced by Polanyi (1966), has been used to characterize the knowledge gained from everyday experience that has an implicit, unarticulated quality (Neisser, 1976; Schön, 1983; Sternberg, 1985a, 1988, 1997a). Sternberg and his colleagues (Sternberg, 1997a, 1997b; Sternberg and Horvath, 1999; Wagner and Sternberg, 1985) view tacit knowledge as an aspect of practical cognition. It is knowledge that reflects the practical skill to learn from experience and to apply that knowledge in pursuit of personally valued goals. Tacit knowledge is needed to successfully adapt to, select, or shape real-world environments. Because tacit knowledge is an aspect of practical cognition, it provides insight into an important factor underlying the successful performance of real-world tasks. Research by Sternberg and his colleagues (see e.g., Sternberg et al., 1993; Sternberg et al., 1995), which we review in later sections of this report, has shown that tacit knowledge can be applied to understanding performance in a variety of job domains.

Support for the importance of the concept of tacit knowledge is found also in research on expertise and implicit learning. Research with experts in a variety of knowledge-intensive domains has shown that reasoning and problem solving in such domains depend upon proceduralized skills and schematicallyorganized knowledge, both of which may operate outside of focal awareness (see Chi, Glaser, and Farr, 1988). Furthermore, expert knowledge appears to reflect the structure of the operating environment or situation more closely than it does the structure of formal, disciplinary knowledge (Groen and Patel, 1988).

Research on implicit learning focuses on the phenomenon of learning without intention or awareness. Tacit knowledge may be, but need not be, acquired implicitly. Arthur Reber and his colleagues' work on the acquisition of stochastic grammars and of event sequences suggested that human beings are capable of acquiring knowledge of a very complex nature without conscious intention or awareness of learning (Reber, 1967, 1969; Reber and Millward, 1968). Researchers subsequently applied the paradigm to study learning of meaningful information (e.g., information about other people and information about the behavior of an economic system) and replicated the basic pattern of results (Broadbent and Aston, 1978; Broadbent, Fitzgerald, and Broadbent, 1986). The research on implicit learning suggests that knowledge can be acquired in the absence of awareness or intention to learn, and thus has a hidden or tacit quality.

In this section, we begin by discussing the type of theoretical concept we consider tacit knowledge to be. Next, we describe the characteristic features of tacit knowledge and how it is distinguished from related concepts. Then, we consider how tacit knowledge

is represented at different levels of abstraction. We present a cognitive model that relates the key features of tacit knowledge to the acquisition, storage, and retrieval of knowledge in and from memory.

3.1 Tacit knowledge as a theoretical concept

In research by Sternberg and his colleagues (Sternberg et al., 1993; Sternberg et al., 1995; Wagner and Sternberg, 1985), the term *tacit knowledge* has been used to characterize a type of knowledge, the possession of which distinguishes more from less practically-successful individuals. In order to understand better the theoretical concept of tacit knowledge, we begin with a distinction between nominal and natural concepts.

Nominal concepts are used attributively. For example, we use the term "bachelor" to attribute certain features (i.e., male, adult, unmarried) to some objects or persons. The instances of a nominal concept often share features that are both necessary (i.e., all valid instances must have these features) and sufficient (i.e., having these features is enough to qualify something as a valid instance). Membership in a nominal concept is "all or none"—either an instance possesses the critical features or it does not.

Natural concepts, in contrast, are used ostensively. For example, we use the term "furniture" to refer to objects that we view as equivalent (e.g., dresser, chair, table). The instances of a natural concept share characteristics features, but these features are not necessary or sufficient for membership. Membership in a natural concept is not "all or none," but rather instances are judged in terms of their strengths of resemblance to the concept. This means that some instances (those with high resemblance) will be judged as better examples of the concept than will other instances (those with low resemblance). For example, most people would agree that "arm chair" is a more typical example of the concept "furniture" than is "bean bag chair."

Tacit knowledge is a natural concept. It is used to denote a type of knowledge that is held together by the resemblance of items to one another and not by a set of individually-necessary and jointly-sufficient features. This lack of necessary and sufficient features does not mean that as a concept tacit knowledge is incoherent or meaningless. Two people may not be able to identify the critical features that all items of furniture share, but they can still agree that furniture exists and that a coffee table is furniture and a toaster oven is not.

Because tacit knowledge is a natural concept, we do not expect that judgments about what is and is not tacit knowledge will be "all or none." Rather judgments should depend on the item's strength of resemblance to the concept. Some knowledge will seem to represent a particularly clear example of tacit knowledge and other knowledge will seem marginal. For marginal items, individuals may disagree about whether the item is a valid instance of tacit knowledge. Given a high level of agreement among judges, the tacit quality of knowledge items can be determined with some degree of confidence.

We describe below three key features that are commonly shared by items of tacit knowledge. These features are used to judge the resemblance of items to the concept. In other words, items that possess these features are more likely to be characteristic of tacit knowledge.

3.2 The characteristic features of tacit knowledge

We identify three key features of tacit knowledge. These features of tacit knowledge relate to (a) the conditions under which it is acquired, (b) its cognitive structure, and

(c) the conditions of its use. First, tacit knowledge generally is acquired on one's own with little support from the environment (e.g., through personal experience rather than through instruction). Second, tacit knowledge is viewed as procedural in nature. It is associated with particular uses in particular situations or classes of situations. Third, because it generally is acquired through one's own experiences, tacit knowledge has practical value to the individual. We expand upon each of these features below.

3.2.1 Tacit knowledge typically is acquired without environmental support

Tacit knowledge generally is acquired on one's own. That is, it is acquired under conditions of minimal environmental support. By environmental support, we mean either people or media that help the individual to acquire the knowledge. As such, tacit knowledge tends to be unspoken, underemphasized, and poorly conveyed relative to its importance for practical success.

When people or media support the acquisition of knowledge, they facilitate three knowledge-acquisition components: selective encoding, selective combination, and selective comparison (Sternberg, 1988). When an individual is helped to distinguish more from less important information (selective encoding), to combine elements of information in useful ways (selective combination), and to identify knowledge in memory that is relevant to the present situation (selective comparison), the individual has been supported in acquiring knowledge. In performing real-world tasks, individuals often must engage in these processes on their own in order to make sense of and respond to situations. The resulting knowledge may reflect the use of these processes, but the individual may not be able to express how the knowledge was acquired.

3.2.2 Tacit knowledge is procedural

The second feature of tacit knowledge is its close association with action. Tacit knowledge takes the form of "knowing how" rather than "knowing that." Anderson (1983) has characterized these two respective types of knowledge as procedural and declarative. More precisely, procedural knowledge is knowledge that is represented in a way that commits it to a particular use or set of uses. It is knowledge that guides behavior, usually without being readily available to conscious introspection. People may not know they possess and/or may find it difficult to articulate such knowledge. We view procedural knowledge as a superset of tacit knowledge. All tacit knowledge is procedural, although not all procedural knowledge is tacit.

The characterization of tacit knowledge as procedural derives from our research. We have found that when individuals are queried about the knowledge they have acquired through their experiences, they often begin by articulating general rules in roughly declarative form (e.g., "a good leader needs to know people"). When these general statements are probed, the statements often reveal themselves to be more abstract or summary representations of a family of complexly specified procedural rules (e.g., rules about how to judge people accurately for a variety of purposes and under a variety of circumstances). These procedural rules, we believe, represent the characteristic structure of tacit knowledge and serves as the basis for identifying and measuring tacit knowledge.

We can represent tacit knowledge in the form of condition-action pairings:

IF <antecedent condition> THEN <consequent action>

For example, the knowledge of how to respond to a red traffic light could be represented as:

IF <light is red> THEN <stop>

Of course, the specification of the conditions and actions that make up proceduralized knowledge may be quite complex. In fact, much of the tacit knowledge that we have observed seems to take the form of complex, multicondition rules (production systems) for how to pursue particular goals in particular situations. In other words, tacit knowledge is more than a set of abstract procedural rules. It is context-specific knowledge about what to do in a given situation or class of situations. For example, knowledge about confronting one's superior might be represented in a form with a compound condition:

IF <you are in a public forum>

AND

IF <the boss says something or does something that you perceive is wrong or inappropriate >

AND

IF <the boss does not ask for questions or comments>

THEN <speak directly to the point of contention and do not make evaluative statements about your boss, staff or your peer's character or motives>

BECAUSE <this saves the boss from embarassment and preserves your relationship with him.>

3.2.3 Tacit knowledge is practically useful

The third characteristic feature of tacit knowledge is its instrumental value in attaining people's personal goals. The more highly valued the goal is, and the more directly the knowledge supports the attainment of the goal, the more useful is the knowledge. For example, knowing that seeking input from subordinates makes them feel valued is practically useful for those supervisors who want their subordinates to feel valued, but not practically useful for supervisors who do not value this goal.

We do not believe that practically useful knowledge must be acquired in any particular context or forum. Useful knowledge is, of course, acquired in classrooms, from experience on the job, through mentoring relationships, and through self-study. We distinguish practically useful knowledge not from formally acquired knowledge but, rather, from knowledge (however acquired) that is not relevant to the practical goals that an individual values.

3.2.4 Tacit knowledge involves coherent relations among its features

The three features of tacit knowledge, acquisition on one's own, procedural structure, and practical value, are related to one another in a non-arbitrary way. That is, we can explain why these features go together in the specification of a meaningful natural concept of tacit knowledge.

First, there is a natural correpondence between the features of procedural structure and practical value. Procedural knowledge tend to be practically useful—it contains within it the specification of how it is to be used. Declarative knowledge, in contrast, is not specific with respect to use and, as a consequence, may remain inert or unused. Therefore, procedural knowledge is more likely to be relevant in the pursuit of personally-valued goals.

Second, knowledge acquired under low environmental support is more likely to have practical value. When knowledge must be acquired on one's own, the probability

increases that some individuals will fail to acquire it. When some individuals fail to acquire knowledge, those who succeed may gain a comparative advantage. This advantage is expected to be lower when the knowledge is highly supported by the environment (i.e., explicity and effectively taught) because more people would be expected to acquire and use it. At the same time, knowledge acquired through one's own experiences should have more personal relevance to the types of situations one encounters in everyday life.

Finally, we associate knowledge acquired through experience with knowledge that is procedural in structure. Because procedural knowledge is more difficult to articulate and more poorly conveyed relative to declarative knowledge, its acquisition is more likely to be a function of experiential learning. By the same token, knowledge acquired through experience is more likely to be related to action because originally it was obtained in the context of performing a practical, everyday task.

Each of these features is viewed as a continuous, rather than discrete, dimension of tacit knowledge. That is, knowledge is not categorized as either possessing or not possessing these features, but rather it is a matter of degree. Some knowledge may be more well-supported by the environment than other knowledge. Similarly, some knowledge may have more practical value to the individual than other knowledge. Knowledge that is closer to one end of the continuum is considered more representative of tacit knowledge.

3.3 What tacit knowledge is not

We have identified above the features that help describe what type of knowledge we consider tacit knowledge to be. It is helpful also to distinguish tacit knowledge conceptually from other related concepts such as job knowledge, general cognition, and performance.

3.3.1 Tacit knowledge is not synonymous with job knowledge

Schmidt and Hunter (1993) suggested that tacit knowledge is merely a type of job knowledge. Tacit knowledge and job knowledge are viewed more appropriately as overlapping concepts. First, some, but not all, tacit knowledge pertains to job-related activities. Tacit knowledge can pertain to any personally-valued activity, including academic and social activities; it is more than job knowledge. Second, some, but not all, job knowledge is tacit. Job knowledge includes declarative and procedural knowledge, with some of the latter characterized as tacit. Job knowledge may be explicit and readily verbalized, as in the rules for operating a lathe or the steps used to compute simple interest, or the knowledge may be tacit, as in knowing what package design will likely sell a product.

Measures of tacit knowledge have the potential to explain individual differences in performance that are not explained by traditional measures of job knowledge, which tend to assess more declarative, explicit forms of knowledge (see e.g., Schmidt and Hunter, 1998). Individual differences in the skill or inclination to acquire and use tacit knowledge make it a potentially useful construct for understanding intelligent behavior in real-world settings, as well as for predicting success in such settings.

3.3.2 Tacit knowledge is not a proxy for general cognition

The skill or propensity to acquire tacit knowledge is viewed as a dimension of practical cognition that conventional skill tests do not adequately measure. Overall cognitive ability tests and similar tests, which are intended to measure so-called general cognition (*g*), are composed of problems that can be characterized as largely academic or abstract.

As discussed earlier, academic problems are well-defined, abstract problems that do not necessarily reflect real-world tasks (Neisser, 1976; Sternberg, 1988, 1997a). Therefore, overall cognitive ability tests and similar tests measure problem-solving skills that are relatively different from the skills needed to solve everyday, practical problems. For this reason, we do not view measures of tacit knowledge as proxies for measures of academic cognition. Although general cognitive skill may support the acquisition and use of tacit knowledge in important ways, tacit knowledge is not reducible to academic cognition. Of course, it is an empirical question whether measures of tacit knowledge do in fact correlate with measures of crystallized cognition. This question is addressed in subsequent sections.

3.3.3 Tacit knowledge is not sufficient for effective performance

Although we do not consider tacit knowledge to be a proxy for general cognition, we do recognize that so-called *g* and other factors contribute to successful performance in many jobs, based on traditional criteria of success (such as performance ratings). The performance of many everyday tasks requires general academic cognition in (at least) the normative range, motivation to succeed, nontacit domain knowledge, and many other resources. We recognize and basically are in concurrence with the results of numerous meta-analyses that show the significant contribution of these variables to understanding performance (see Schmidt and Hunter, 1998). But we attempt to supplement these variables and improve upon conventional approaches to understanding, predicting, and improving performance in real-world settings.

Measures of practical cognition, like all measures of cognition, are, at best, indicators of the underlying cognitive functions we seek to understand. As such, we can talk about practical cognition, and more specifically tacit knowledge, at different levels of abstraction. That is, we can conceptualize tacit knowledge at the level of its cognitive representation, and at the level which it is measured in the behavior and articulated knowledge of the individual. We discuss these different levels of abstraction below.

3.4 Describing tacit knowledge at different levels of abstraction

Tacit knowledge can be conceptualized at qualitatively different levels of abstraction. At the lowest, least abstract level, tacit knowledge can be described as mentally-represented knowledge structures. We believe that these knowledge structures take the form of complex, condition-action mappings. At this level of description, tacit knowledge takes on its psychological reality and has its consequences for intelligent behavior.

Ideally, we would measure the possession of tacit knowledge directly at the level of its cognitive representation. However, we must infer possession of tacit knowledge from the knowledge that people articulate. When knowledge is articulated, often it is greatly simplified. That is, the complex knowledge structures that map sets of antecedent conditions onto consequent actions are summarized and abbreviated into general rules and procedures. It is at this level, that we measure people's tacit knowledge.

At a higher, more abstract level of description, tacit-knowledge items can be grouped into categories of functionally-related items. Describing tacit knowledge at this level adds value to the identification of tacit knowledge by highlighting the broad, functional areas or competencies that tacit knowledge represents. In other words, in addition to specific items of tacit knowledge, we can identify more generally the types of knowledge that are likely to be tacit.

3.4.1 Identifying and measuring tacit knowledge

Measuring tacit knowledge takes into account the realistic, contextualized quality of the knowledge. Responses to realistic problem situations are used as indicators of an individual's possession of tacit knowledge. Wagner and Sternberg (1985) devised a method of presenting scenarios to individuals that depict the types of problems they face in their given pursuits. These scenarios reflect the types of situations in which recognized domain experts have acquired knowledge characterized as "tacit." Because tacit knowledge is not readily articulated, we rely on observable indicators (e.g., responses to the scenarios) to assess whether an individual possesses knowledge characterized as tacit, and can apply that knowledge to the situation at hand. The responses reflect an individual's skill to recognize and take appropriate action in a given situation, and presumably, their procedural knowledge.

Deriving the information for these scenarios poses a challenge in that the tacit knowledge of domain experts must somehow be identified. Domain experts are appropriate sources for identifying tacit knowledge because in order to achieve their expert status, they likely have acquired knowledge that others have not (i.e., knowledge without direct support). As a subset of procedural knowledge that is not readily articulated, tacit knowledge is not likely to be elicited directly from individuals. However, since tacit knowledge is experience-based, we attempt to identify the knowledge in the recalled experiences of individuals. In other words, when individuals have difficulty expressing their action-oriented knowledge, we attempt to elicit memories for the particular episodes that produced that knowledge.

In the next section, we describe methods used to elicit examples of tacit knowledge from domain experts and to develop instruments to measure the acquisition and use of tacit knowledge within a given domain. The methods, which have been applied in domains ranging from education to military leadership, have evolved over the course of our tacit-knowledge research, resulting in a refined and detailed methodology for eliciting and measuring tacit knowledge. We devote the next section to describing this methodology as it plays an important role in understanding the findings from tacit-knowledge research and offers a tool for studying tacit knowledge in any domain.

4. Measuring tacit knowledge

One of the goals of our research is to show that tacit knowledge contributes to successful performance in a variety of domains. That is, we aim to establish a relationship between the possession of tacit knowledge and performance. But how does one proceed to develop a test to measure tacit knowledge? This section addresses the development of tools to measure the amount of tacit knowledge of various kinds that an individual has acquired. We begin by reviewing some approaches that have been used to measure the competencies considered to be relevant to the performance of real-world tasks, and contrast them with our knowledge-based approach. We then discuss what tacit-knowledge tests are intended to measure and offer a general framework for developing and validating such a test through the assessment of everyday situational judgments.

4.1 Methods of measuring real-world competencies

The tacit-knowledge approach to understanding practical cognition is based on several methods of measuring real-world competencies. These include the use of the critical-incident technique, simulations, and situational-judgement tests. We review briefly each of these methods and then discuss how the tacit-knowledge approach draws certain aspects from these methods.

4.1.1 Critical-incident technique

The critical-incident technique is an approach that seeks to identify the behaviors associated with effective performance (Flanagan, 1954). According to Flanagan, a critical incident describes the behavior, the setting in which the behavior occurred, and the consequences of the behavior. Critical incidents are generated by asking individuals, typically subject-matter experts, to provide examples of effective and ineffective behaviors. More specifically, individuals are asked, through interviews or open-ended survey questions, to describe several incidents that they, or someone else, handled particularly well, as well as several incidents that they, or someone else, handled poorly (Flanagan, 1954; McClelland, 1976). Boyatzis (1982) used a variation on the critical-incident technique, called the "behavioral event interview," in which he obtained behavioral incidents from individuals identified a priori as either high, medium, or low on effectiveness. He then examined the incidents generated from each group to identify traits and skills that distinguished between effective and ineffective managers.

The "critical incidents" generated from observations, interviews, or surveys are analyzed qualitatively to determine the nature of the competencies that appear important for success in a given task domain. The incidents typically are grouped on the basis of similar behavior content. For example, an incident that pertains to assigning a task to a subordinate and an incident about monitoring task completion by a subordinate might be grouped into a category of supervising subordinates. These categories are used to draw general conclusions about the behaviors that are characteristic of effective and ineffective performers.

Limitations of the critical-incident technique are that it assumes people can and will provide incidents that are critical to success in their particular jobs, and that qualitative analysis is sufficient for identifying the underlying competencies. However, the value of the critical-incident technique lies in identifying the strategies individuals use to perform various tasks, and in examining specific, situationally-relevant aspects of behavior. The critical-incident technique has been used successfully in the development of several performance assessment tools, including behaviorally anchored rating scales (BARS; e.g., Smith and Kendall, 1963) and situational-judgment tests (SJTs; e.g.,

Motowidlo, Dunnette, and Carter, 1990), the latter of which is described in more detail below.

4.1.2 Simulations

Simulations have been used as both assessment tools and as training methods. Simulations are aimed at assessing job behaviors directly. They involve observing people in situations that have been created to simulate aspects of the actual job situation. Responses to these simulations are considered to represent the actual responses that individuals would exhibit in real situations. Simulations can take the form of in-basket tests, situational interviews, group discussion, assessment centers, and situational-judgement tests. Motowidlo et al. (1990) distinguished between high-fidelity and low-fidelity simulations. In high-fidelity simulations, the stimuli presented to the respondent closely replicate the actual situation and the individual has an opportunity to respond as if they were in the actual situation. In low-fidelity simulations, the stimuli are presented in written or oral form and the individual is asked to describe how he or she would respond to the situation, rather than actually to carry out the behavior.

At the high-fidelity end of the continuum is the assessment center. Assessment centers present small groups of individuals with a variety of tasks, including in-basket tests, simulated interviews, and simulated group discussions (Bray, 1982; Thornton and Byham, 1982). The simulation approach has the advantage of more closely representing actual job performance. However, it is not always clear what aspects of the job should be chosen to simulate or how to evaluate performance.

In-basket tests have a moderate level of fidelity. In an in-basket test, the participant is presented with various materials (e.g., memos, financial reports, letters) and is asked to respond to them (Frederiksen, 1966; Frederiksen, Saunders, and Wand, 1957). The individual, however, has a limited amount of time to deal with the problems presented in the in-basket, giving him or her some of the constraints of actual job situations. Performance is evaluated based on how the items are handled. For example, does the participant respond to a letter from the Director of Finance requesting fourth-quarter financial records with complete and accurate information?

Situational-judgment tests have been considered low-fidelity simulations (see Motowidlo et al., 1990). Situational-judgment tests (SJTs) present descriptions of situations, typically work-related, in which a problem exists (see e.g., Chan and Schmitt, 1998; Legree, 1995; Motowidlo et al., 1990). The descriptions can be of actual situations or written to approximate actual situations in the domain of interest (e.g., a salesperson making a phone solicitation). Situations typically are selected on the bases of a critical-incident analysis. Following each situational description is a set of options (i.e., strategies) for solving the problem. Respondents are asked to indicate their endorsement of the options, either by selecting the best and possibly the worst from among a few strategies, or rating the effectiveness of each alternative. Traditionally, SJTs have been scored by awarding points based on the correct choice of the best and worst options (e.g., Motowidlo et al., 1990), or awarding points based on the percentage of experts who endorse the option (e.g., Chan and Schmitt, 1998).

4.1.3 Tacit-knowledge approach

The tacit-knowledge approach draws on aspects of the above approaches in order to measure the level of domain-specific, procedural knowledge that individuals have acquired from solving everyday problems. It is based on theoretical and empirical claims that the amount and organization of knowledge that experts possess differs from that of novices (see Chi et al., 1988; Ericsson, 1996; Ericsson and Smith, 1991), and that these

knowledge differences reflect differences in the developed skills of experts and novices (Sternberg, 1998a; in press-a).

The tacit-knowledge approach relies on a critical-incident technique to identify examples of tacit knowledge acquired in solving real-word problems. That is, we interview domain experts to identify incidents that reflect important learning lessons, and ask them to express in their own words the knowledge gained from those situations. We do not rely solely on the individuals who provided the incidents to determine which items of knowledge are more or less effective. We use subsequent analyses to identify the items that are "critical" to performance.

The tacit-knowledge approach shares with the simulation approach the view that measuring practically relevant behavior in a test situation depends, in part, on the extent to which the task resembles those tasks found in everyday life. As such, we attempt to include sufficient detail in our measure to provide respondents with a realistic picture of the situation. However, we have relied primarily on a paper-and-pencil format to present this information rather than simulations for reasons of practicality, with the exception of our tacit-knowledge-acquisition task for sales (Sternberg et al., 1993). We have chosen to provide better coverage of the performance domain at the potential cost of lower fidelity. Future testing, however, is moving in the direction of more performance-based, high-fidelity assessment.

The tacit-knowledge approach is linked most closely to that of situational-judgment testing. We present situation descriptions, often based on actual situations of position incumbents, followed by several possible responses to those situations. The number of response options range between five and twenty. Individuals are asked to rate on a Likert scale the quality or appropriateness of each option for addressing the problem presented in the situation.

For example, in a hypothetical situation, an administrative assistant realizes that there is a factual error in a memo her boss has written and the memo needs to be sent out immediately. The boss is in a closed-door meeting. The respondent is asked to rate several options (usually on a 1 = low to 9 = high scale) for solving the problem. Examples of responses include (a) interrupting the meeting to show the boss the error, (b) fixing the error oneself and sending out the revision, and (c) fixing the error but waiting to send out the memo until the assistant can run it by the boss.

The set of ratings the individual generates for all the situations is used to assess the individual's tacit knowledge for that domain. Similar to SJTs, the scoring of tacit-knowledge tests often rely on the judgments of experts. In general, tacit-knowledge tests have been scored in one of three ways: (a) by correlating participants' responses with an index of group membership (i.e., expert, intermediate, novice), (b) by judging the degree to which participants' responses conform to professional "rules of thumb," or (c) by computing the difference between participants' responses and an expert prototype. To understand better what tacit-knowledge tests are designed to measure, we consider tacit knowledge as a measurement construct.

4.2 Tacit knowledge as a measurement construct

What are tacit-knowledge tests, and the items contained within them, intended to measure?

This question can be answered by considering a traditional distinction between achievement testing and cognition testing. In achievement testing, items are presumed to exemplify the measurement construct (e.g., knowledge of world history) but are not commonly viewed as predictors. For example, when an individual correctly answers a

factual, multiple-choice question about world history, we assume that she possessed prior knowledge of either the fact in question or related facts that enabled her to rule out incorrect alternatives. We do not commonly view the history question as predictive of performance on other tests or tasks. In cognition testing, by contrast, items are presumed to predict performance but are not commonly viewed as exemplars of the measurement construct. For example, when an individual correctly solves a figural analogy problem, we do not assume that he possessed prior knowledge of the analogical relationship in question. However, we do view such analogy problems as predictive of performance on other tests and tasks of general mental skill.

Is a measure of tacit knowledge a cognition test or an achievement test? Having drawn a distinction between cognition and achievement testing, we must point out that neither type of test exists in a pure form (Sternberg, 1998a). All achievement tests measure underlying skills—if only the skills necessary to acquire and display mastery of the tested content—and so tend to have predictive value. Likewise, all cognition tests measure acculturated knowledge—if only the knowledge necessary to make sense of items and testing conventions—and so tell us something about the knowledge content of individuals rated high and low in general cognition. All of these tests measure a form of developing expertise (Sternberg, 1998a). Tacit-knowledge tests break down the (artificial) boundaries between achievement and skill testing.

Tacit-knowledge tests are everyday situational-judgment tests built on a theory of human cognition (Sternberg, 1995c). They are intended to measure both practical, experience-based knowledge and the underlying dispositions or skills that support the acquisition and use of that knowledge. Thus, scores on tacit-knowledge tests are expected to predict performance on tests or tasks that draw on either tacit knowledge or the mental skills that supported its development and use. These skills are hypothesized to differ from those implicated in the so-called "general factor" in human cognition commonly referred to as *g* and often approximately measured, in norm-referenced fashion, as IQ. Research by Sternberg and colleagues has produced support for the hypothesis that the skills associated with tacit-knowledge test performance are different than those associated with tests of *g* (Hedlund et al., 1999; Sternberg et al., 1993; Sternberg et al., 1995).

Because tacit-knowledge items are considered to measure both acquired knowledge and practical skill, we propose that tacit-knowledge tests have the potential to shed light upon (1) the content of tacit knowledge and (2) the events or experiences through which it was acquired. Few would contest that tacit-knowledge items reflect the knowledge of the respondents from whom the items were obtained (in the course of a "story-telling" exercise focusing on personal experiences). The items came from these respondents' memories and so must reflect the content of those memories. What remains to be determined is the degree to which tacit-knowledge items measure the acquisition and use of tacit knowledge by those who did not produce but, rather, endorsed or rated the items. This question is addressed by our numerous research studies in both civilian and military sectors, which we discuss in subsequent sections.

4.3 Developing everyday situational-judgment tests to measure tacit knowledge

We have developed tests to assess tacit knowledge for academic psychology, elementary-school teaching, business management, sales, entry-level jobs in organizations, college education, and military leadership. In this section we present a framework for developing tacit-knowledge tests of the format described above, a framework that is based on the techniques we have used to measure tacit knowledge in the various domains we have studied.

The development of tacit-knowledge inventories readily may be understood as a production process, beginning with the "raw materials" of experience-based tacit knowledge elicited from successful practitioners in a given domain and culminating in a revised and validated inventory. At each step in the development process, "value" is added through the conduct of research and analysis.

All of the phases are designed to support the development of assessment instruments based on (a) the theory and methods of tacit-knowledge research, and (b) the substantive knowledge in the domain of interest. Specifically, the steps are intended to aid in selecting the content that is most promising with respect to the goals of the assessment phase, that is, in measuring an individual's possession of tacit knowledge. The term *promising* is used here to refer to that subset of tacit knowledge with the highest probability of yielding or contributing to tacit-knowledge test questions that, taken together, constitute a valid measure of the underlying, domain-relevant tacit knowledge of respondents. This process was developed over the course of several research projects, and is applicable to the identification and assessment in tacit knowledge in any performance domain. We describe each stage in the process below, from the identification of exemplars of tacit knowledge to the construction of the final inventory.

4.3.1 Knowledge identification

We generally begin with a review of job-relevant literature (e.g., sales manuals, Army trade publications) to identify on a preliminary basis the experience-based, tacit knowledge for the relevant profession (e.g., salespersons, Army leaders). This review may suggest some of the content for use in a tacit-knowledge inventory, and may provide a preliminary taxonomy, or category framework, for organizing the knowledge. For example, in research with managers, Wagner and Sternberg (1986) proposed a framework of practically-intelligent behavior consisting of tacit knowledge about managing oneself, managing others, and managing one's career.

Typically, a review of the literature does not provide a sufficient number of examples of knowledge that meet our criteria or include enough detail from which to create tacit-knowledge questions of the format described above. We have found that the practical advice presented in the professional literature tends to be decontextualized and already converted to semantic knowledge. We also surmise that the politics of professional print may keep some truly tacit knowledge—knowledge that contradicts doctrine, for example—out of print altogether. Therefore, the next step is to conduct interviews with successful practitioners in the domain to generate a larger body of knowledge from which to draw in developing the tacit-knowledge inventories. We described here a method for conducting these interviews.

A method for eliciting tacit knowledge. In selecting individuals to interview, it is important to identify a sample that is likely to possess a certain amount of tacit knowledge. We seek to identify individuals who are both experienced and successful in their domain. Individuals who are more successful likely have acquired some important knowledge relevant to success that individuals who are less successful have not. Furthermore, individuals who are currently practicing in the domain of interest are more appropriate sources for understanding the tacit knowledge of that domain than are individuals who hold other positions (e.g., supervisor) or previously held the position of interest. The latter may consider different knowledge to be relevant, based on their different perspectives. Once a relevant pool of practitioners is identified, experts can be chosen either through nominations (e.g., by peers or superiors) or based on existing performance criteria (e.g., performance evaluation, salary). In research by Sternberg and his colleagues (Hedlund et al., 1999; Sternberg et al., 1993; Sternberg et al., 1995; Wagner, 1987), interviews were conducted with academic psychologists deemed successful based on

their tenure and affiliation (e.g., full professors at Yale); business managers who were considered successful on the basis of their position in the company; salespersons who were successful in their sales performance; successful college students selected based on grades and school affiliation; and successful military leaders identified through a nomination process.

All of these experts were asked to consider what it takes to succeed in their respective domains and to provide typical performance-related situations and possible responses to those situations that exemplify tacit knowledge. In recent research, we have developed a structured interview in which participants are provided with more explicit instructions about the knowledge we seek to identify and which prompts them to elicit more in-depth responses.

We rely generally on a two-person interview team, with one person designated as the lead interviewer and the other the notetaker. The lead interviewer directs the interview and the notetaker takes written notes, asks for clarification, and, along with the lead interviewer, asks follow-up questions. The interviews also are taped when possible, with the interviewees' consent, so that questions and clarifications can be addressed once the interview is completed. It is also helpful for one of the interviewers to be familiar with the domain in order to understand any technical language or the jargon of the interviewee.

We present below a protocol for conducting the interviews. We use specific examples from our work with military leaders to illustrate the steps involved.

1. Introduction. When the participant arrives, members of the interview team introduce themselves and give a standardized introduction to the study and the interview. This introduction should state the purpose of the research, preempt likely misunderstandings, and orient the participant to the purpose of the interview. For example:

   We are trying to understand the key lessons that leaders acquire from their experience on the job. If we can identify these lessons, we will try to find ways to use them to strengthen leader development efforts.

   This is not an evaluation of you as a leader. This is not a study comparing leaders from your organization to those from another organization.

   We want to identify specific examples of informal knowledge about leadership at your level. We want to find examples of things about leadership that are not written in books or taught in classes. Our belief is that this knowledge is often not discussed openly, but nevertheless is used by leaders as they meet the demands of their jobs. This knowledge may have been learned because of some challenge or problem you faced. It may have been acquired by watching someone else's successes or failures.

   We are not interested in the "party line" or "doctrine." We also are not interested in the purely technical things you learned from experience (e.g., how to tune up an engine). We are really interested in the problems and challenges you face and what you have learned about leadership at your level from these experiences.

2. Request for stories. The purpose of the interview is to elicit stories or cases from the participants' experiences and to explore the unspoken, practical knowledge gained from or reflected in those cases. We ask participants, for example, to:

   "Tell us a story about a leadership experience you have had as a leader in your current position from which you learned a lesson."

The aim is to keep the focus on the stories rather than theories or generalizations about effective performance. In this way, the responses are more closely tied to the tacit-knowledge construct (i.e., in the knowledge based on personal, practical experience).

Because the values in the recalled experiences are sometimes unclear, we seek the participant's help in making sense of each story, and identifying the lesson associated with the story.

3. Follow-up questions. Follow-up questions are used to focus on key contextual variables in the stories (e.g., "Tell us more about the climate in your unit"); the goals and alternative courses of action reflected in the stories (e.g., "What exactly did you hope to accomplish?" and "What else did you consider doing at the time?"); and on identifying practical knowledge with broader applicability (i.e., "lessons learned") derived from the experiences described in the stories (e.g., "What do you think you learned from this experience?" and "How has this experience affected your approach to leadership?"). Once it appears that no more information can be gained from a story, the interviewer, given time allowances, may ask the participant to share another story from his or her experience.

At the completion of each interview, the notetaker summarizes the interview. An interview summary might contain the following information: (1) participant information (e.g., position, time in job, race, gender), (2) a summary of each story, (3) annotations to each story based on follow-up questions, and (4) any comments from the interviewer. It is useful for the notetaker and lead interviewer to review the summaries and resolve any disagreements over details or interpretations from the interview.

The identification of tacit knowledge does not end with the summarized interviews. Even with explicit instructions about what the interviewer is looking for, not all of the stories generated from the interviews provide examples of tacit knowledge. Therefore, the interview summaries are submitted to a panel of experts who are familiar with both the performance domain and the tacit-knowledge construct. These experts are asked to judge whether the interview summary represents knowledge that is intimately related to action, is relevant to the goals that the individual values, is acquired with minimal environmental support, and is relevant to performance in the domain under study (e.g., academic psychology, military leadership).

Products of the interviews. The products of the interviews are transcripts and summaries that contain numerous potential examples of tacit knowledge. These summaries serve two purposes in instrument development. First, tacit-knowledge "items" (essentially pieces of advice) may be extracted from the summaries and used in a number of later analyses. Second, the summaries themselves (consisting of stories that the professionals shared about their experiences) can be used directly in the construction of the inventory.

A useful interim step is to ask a panel of experts (e.g., members of the research team or practitioners familiar with the tacit-knowledge construct) to review the knowledge compiled from the interview summaries to ensure that it meets the criteria for tacitness. These criteria are that (1) the knowledge should have been acquired with little environmental support, (2) it should be related to action, and (3) it should have relevance to the goals that the person values. Often, upon further review, a knowledge example may be judged by experts to fail to meet one of these criteria. For example, consider the following story told by a military officer.

> *I had a lieutenant who was screwing up big-time. He would take sensitive items (e.g., weapons, night-vision devices, etc.) home. He even lost sensitive items. He lost a pistol, and rather than stop the mission and look for it, he continued on with the mission. As we all know, when you lose a sensitive item, you stop everything and look for it until you find it.*

The above story was deemed to lack the necessary criteria for tacitness. The interviewee indicated that the knowledge he referred to is generally known by leaders. It even may represent an official procedure. Therefore, we have no evidence that this knowledge is attributable to the officer's experience in dealing with sensitive items that are missing. On the other hand, consider a story from another officer about a similar issue.

> *It is important for a commander to know when to report bad news to the boss and when to withhold it. My unit had just completed a night move and had been in position for about two hours. A weapon was identified as missing around midnight. The section chief told me that the weapon was in the current position because he had seen it during the sensitive item checks. I talked to each member of the section and determined that the weapon was in the position. We looked for the weapon from about midnight until 0300 hours. During this time I chose not to notify the battalion commander because I was confident that the weapon would be found. However, a sensitive item report was due at 0400 hours, so, for ethical reasons, I notified the battalion commander at 0300 hours that the weapon was missing. I told the battalion commander what I had done so far and that I was confident that the weapon would be found at first light. He was not upset. We found the weapon within ten minutes after the sun came up and the battalion commander was pleased we followed the standard operating procedures for dealing with a missing weapon.*

In this story, the officer clearly expresses some knowledge he has acquired through previous experience in dealing with missing sensitive items (e.g., weapons). He has learned that, under some circumstances, it is best to hold off reporting a problem until it becomes necessary, so long as appropriate steps are taken to resolve the problem in the interim.

Coding the interview summaries. After determining which examples of knowledge meet the established criteria, it is useful to transform the summaries into a more usable form for the purpose of later analyses. We have used a format that is based on the procedural feature of our definition of tacit knowledge. That is, the knowledge is expressed as a mapping between a set of antecedent conditions and a set of consequent actions.

An item of knowledge is represented by one or more antecedent condition or "IF" statements, by one or more consequent action or "THEN" statements, and by a brief explanation or "BECAUSE" statement. The logical operators "AND" and "OR" are used in the coding to signal relationships of conjunction and disjunction, respectively. The operator "ELSE" is employed in the coding to connect sets of condition-action mappings into more complex procedures. Each individual piece of tacit knowledge is rewritten into this procedural form. This coding allows the researcher to analyze more readily the content of the tacit knowledge for the purpose of identifying categories of knowledge and selecting examples of knowledge that may be useful as items in a tacit-knowledge inventory. The result of this phase is a set of coded tacit-knowledge items.

The coded tacit-knowledge items then may be subjected to a sorting process to identify major categories of tacit knowledge. This sorting may entail asking a group of experts to organize the items according to categories of their own devising. The results of the independent sortings may be analyzed using hierarchical or other cluster analyses, a family of techniques for uncovering the natural groupings in a set of data (for more details regarding this technique, see Hartigan, 1975). This type of analysis may produce

hierarchically organized clusters of items that can be expressed in the form of a tree. The clusters can be interpreted by experts and assigned labels that represent different categories of tacit knowledge. The categories may provide an indication of the major areas of learning that occur in one's respective field. The category framework is also useful in selecting items for test development that provide a broad representation of the performance domain.

4.3.2 Item selection

Although one may proceed to develop test questions directly from the tacit-knowledge items generated from the interviews, a further selection process may be necessary for a number of reasons. First, the interview study may yield too many items of tacit knowledge to include in a tacit-knowledge inventory of reasonable length, depending on the context in which the test might be used. Second, we cannot determine on the basis of the interviews alone what tacit knowledge is diagnostic of experience or predictive of effective performance in a given domain, or alternatively, what tacit knowledge is not related to these criteria. A manager, for example, may have learned that subordinates are more likely to come to her with problems if she leaves her door open. But the extent to which this practice contributes to her success is unclear. By leaving her door open she may become the repository for problems that are the responsibility of other managers, which may create a distraction for her from her job. Third, the results of the preliminary sorting of interview data may not be sufficient for determining the internal structure of the tacit-knowledge construct domain. That is, for the purposes of test construction, we would want further evidence of the structure of the performance domain to ensure the representativeness of our items. For the reasons above, we take an additional step to narrow down the pool of items from which test questions will be constructed.

The next step in the process of selecting items for instrument development is more quantitative than qualitative. It entails surveying job incumbents to assess the "quality" of each tacit-knowledge item. In order to develop a questionnaire that can be administered to job incumbents, the tacit-knowledge items may need to be condensed. For example, if we want professionals to evaluate 100 examples of tacit knowledge, it would be unreasonable to ask them to read 100 items in a very long and complex format. Therefore, it may become necessary to condense the items into briefer descriptions. Condensing the items involves extracting only the most important information and deleting unnecessary information. Attempts should be made to increase the comprehensibility of the items for the intended audience and to preserve the intent of the interviewee who provided the knowledge. The procedural structure that we consider to be characteristic of tacit knowledge is maintained in the rewriting of items.

The condensed items are compiled into a survey, which we refer to as a Tacit Knowledge Survey (TKS), which is in turn a situational-judgment test. A TKS differs from a tacit-knowledge inventory in that respondents are asked to rate the perceived quality of the tacit knowledge in the former, whereas they are asked to rate the quality of responses to the problem in the latter. Job incumbents can be asked to rate each item on a number of dimensions. We have used four seven-point scales that ask for the following judgments: (1) how *good* does the respondent think the advice is, (2) how commonly *known* does the respondent think the advice is, (3) how *often*, in the judgment of the respondent, do incumbents at the specified level face situations such as the one described, and (4) to what extent does the advice match the respondent's personal *concept* of job performance? Each of the scales is intended to provide a different sort of information about the tacit-knowledge item being rated. The "good" scale is intended to assess the overall quality of the knowledge being rated. The "known" scale is intended to assess one possible index of tacitness (i.e., on the theory that knowledge whose

acquisition is not well supported by the environment may be less commonly known than other knowledge). The "often" scale is intended to assess the generalizability or applicability of knowledge items across job settings within the domain. Finally, the "concept" scale is intended to assess respondents' implicit theories of performance. Together, the four rating scales are intended to provide a comprehensive but non-redundant picture of each tacit-knowledge item for the purpose of evaluating each item's potential for development into tacit-knowledge test questions.

We are interested in items that are (1) rated as better advice by those considered to be successful in their domain; (2) not viewed as common knowledge by individuals in the domain; (3) representative of the situations faced by most individuals in the domain; and (4) a good fit to the concept of performance held by successful individuals in the domain. In order to identify items that are endorsed by individuals who are successful in a domain, we obtain data on a relevant performance criterion. In our research with military leaders, we obtained two criterion measures—experience and performance ratings. Experience was expressed in terms of expert-novice differences and performance was assessed using ratings of leadership effectiveness by other leaders. Responses to the TKS are analyzed along with the criterion measure to identify items that have promise for inclusion in the tacit-knowledge inventory. This analysis generates a number of item statistics that can be used in the selection process.

In our research, we used discriminant analysis to identify items that distinguish individuals with more from those with less experience (see Hedlund et al., 1999). In the discriminant analysis, a linear combination of the discriminating variables (e.g., TKS items) is derived that maximizes the divergence between groups (e.g., experienced/novice). The linear combination of the discriminating variables (the canonical discriminant function or CDF) can be tested for significance to determine if the set of variables distinguishes between groups. In addition, the correlations between discriminating variables and the CDF can be computed to assess the discriminating power of individual variables (e.g., TKS items).

We used point-biserial correlations between ratings on the items and ratings of effective performance to identify items that reflected the responses of effective performers. Item statistics such as these can be used, along with the category framework developed in the interview phase, to select items that have the most potential to explain successful performance and provide the best "coverage" of the tacit-knowledge domain.

4.3.3 Instrument construction

The "knowledge identification" and "item selection" phases generate several outputs that serve as materials for the final phase of "instrument construction." These outputs include: (a) interview transcripts and interview summaries, (b) the category framework derived from expert sortings and cluster analyses, (c) a set of item statistics for use in the selection of content for the inventories, and (d) the knowledge items retained on the basis of the category framework and item statistics from the questionnaire study. In the next phase of test development, preliminary inventory questions are constructed, using both selected knowledge items and the interview summaries from which they were drawn. A tacit-knowledge question consists of a situation description followed by several potential responses to that situation. Although the condensed tacit-knowledge item may serve to describe the situation, it is preferable to include the details from the original story to provide a richer, more in-depth problem description. Including more contextual and situation-specific information in the question provides the respondent with a clearer basis on which to evaluate the appropriateness of potential responses to the situation. The original story also provides a source for developing the response options to a question.

Once the researchers are satisfied with the form of the preliminary inventory, it is useful to circulate the inventory among experts in the domain. One method of obtaining feedback is to convene a focus group of experts to review and discuss the inventory. In our research, focus-group participants were given a brief introduction to the goals of the project and an explanation of the tacit-knowledge construct in non-technical language. They were asked to judge the construct-relatedness of the inventory questions by considering whether each question addresses knowledge gained through experience and fits the definition of tacit knowledge provided. In addition, focus group participants were asked to help "fill gaps" and "fix problems" in the inventory. In particular, they were asked to (a) provide additional, plausible response options for any question; (b) identify areas of confusion or lack of clarity; (c) identify problems of gender, racial, or ethnic bias; and (d) identify anything that did not "ring true" in the inventory questions.

The researcher can use the feedback from the focus group to revise the inventories. For example, inventory questions for which judgments of construct-relatedness are not unanimous (and positive) may be omitted from the inventory. Similarly, a response option or scenario feature that is objected to by two or more participants may be omitted. The focus group may suggest additional response options or scenario features, which can be added to the inventory. The final result of this test-development process is a revised tacit-knowledge inventory that can be administered to position incumbents and used to address further research questions, such as those regarding criterion-related construct validity.

4.3.4 Summary

The phases described above all are designed to support the construction of tacit-knowledge tests. The tacit-knowledge items acquired in the interview study form the raw materials for this construction process. During this process, the tacit-knowledge items are subjected to qualitative analysis (e.g., sorting into categories) and quantitative analysis (e.g., obtaining quality ratings). The various phases serve to address two basic questions about the pool of tacit-knowledge from which an instrument will be developed. First, which items are most promising for use in the construction of tacit-knowledge test questions? Second, what does the underlying structure represented by the tacit-knowledge items tell us about the structure of the construct domain so that we can design our tacit-knowledge tests to capture this domain? The result of this process is an inventory that has greater likelihood of possessing both internal and external validity. We discuss the issue of validity in the last part of this section.

4.4 Establishing the validity of tacit-knowledge inventories

An important part of developing any tests is to establish its construct validity. Unlike the development of many cognition-type tests, we do not rely solely on the qualifications that items should load heavily on a single factor and predict some external performance criteria as sufficient for concluding that a test measures the construct of interest. As Nunally (1970) and others have argued, such a "criterion-based" approach to test development is problematic and often produces measurement instruments of inferior quality. Specifically, such an approach may yield tests that suffer from low internal-consistency reliability, poor factor structure, and fragility with respect to criteria other than those on which the selection of items was based.

We rely on both theoretical and empirical justifications to establish the validity of tacit-knowledge tests. We use Messick's (1995) unified validity framework to show how tacit-knowledge theory and the phases of test development outlined above,

contribute to the validity of our tacit-knowledge tests. Messick's framework treats the traditionally separate forms of validity (i.e., content, construct, and criterion) as aspects of a more comprehensive kind of construct validity. According to this framework, the essential goal of test validation is to support, through a combination of theoretical rationale and empirical evidence, the interpretation of test scores and the uses of scores under that interpretation.

4.4.1 The content aspect

The content aspect of validity refers to evidence that test content is relevant to and representative of the focal construct. It addresses the concerns that fall under the traditional heading of content validity. In the context of tacit-knowledge test development, the goal of *construct relevance* calls for tacit-knowledge test questions that are sensitive to knowledge of the type specified by the focal construct and insensitive to knowledge that falls outside the focal construct. A first step toward this goal is taken during the identification phase of test development, in interviews with job incumbents, when we orient participants toward personal experiences and away from formal principles or theory within their performance domains. A second step is taken in the item-selection phase when incumbents are asked to rate the quality of tacit-knowledge items. These ratings (i.e., item means and variances) may provide evidence regarding the relevance of tacit-knowledge items to the underlying construct. For example, tacit-knowledge items with low mean ratings (i.e., when respondents, on average, consider the knowledge represented in the item to be bad advice) may not be relevant to successful performance. And items with low variances (i.e., when respondents agree highly about the quality—good or bad—of the knowledge reflected in the item) may not reflect knowledge gained through personal experience if the knowledge is generally agreed upon as good. In addition to these steps, the goal of establishing construct relevance also is supported by asking domain experts, at various stages in the test development process, to judge the relevance of the items to the tacit-knowledge construct.

The goal of *construct representativeness* calls for tacit-knowledge items that are typical rather than atypical of knowledge-based items specified by the focal construct. An initial step toward this goal is taken in the identification phase by interviewing job incumbents that are representative of the range of specialty areas within the domain. For example, military leaders in the same position (e.g., platoon leader) may serve in one of many branches (e.g., infantry, engineering). Therefore, in our research, we sought to interview officers from these various branches to increase the representativeness of the knowledge that was elicited. A second step is taken during the item-selection phase, when participants are asked to rate how "often" a situation presented in a tacit-knowledge item occurs. Items that receive both a low mean and small variance, for example, are ones that most incumbents agree occur almost never, and therefore may not be representative of the knowledge domain. The categories derived from cluster analyses of the tacit-knowledge items also provide a source for ensuring construct representativeness. Items can be chosen to represent each of the major categories of tacit knowledge, thus providing better coverage of the construct domain. Finally, at several points during test development, expert judgments are sought regarding the construct representativeness of the items. After an initial pool of potential tacit-knowledge items is obtained from the interviews, an expert panel is asked to judge the representativeness of each item. The experts are asked to eliminate items that are too narrow or technical in focus (e.g., how to safely store chemical weapons) and knowledge that is relevant to a small proportion of job incumbents (e.g., how to manage stress at work if you are a single mom). Experts again are asked to evaluate the representativeness of the items after preliminary test questions have been developed.

4.4.2 The substantive aspect

The substantive aspect of validity refers to the theoretical rationale behind tacit knowledge and its relationship to task (test) performance. A step toward the goal of substantive validity is provided by our cognitive model and the characterization of tacit knowledge presented in the sixth section. The model illustrates how tacit, procedural knowledge is acquired and how it comes to be applied in solving everyday problems. The model also helps to illustrate how tacit knowledge confers a performance advantage (relative to that conferred by nontacit, procedural knowledge) in people's skill to respond to contextualized problems of realistic complexity. The characteristic features of tacit knowledge (i.e., acquisition on one's own, procedural nature, and instrumental value) further highlight its potential contribution to successful performance. The cognitive model of tacit knowledge, on which the identification and measurement of tacit knowledge is based, provides a theoretical rationale for tacit-knowledge test performance and, as such, directly serves the goal of substantive validity. Substantive validity also may be supported by showing, through empirical research, the extent to which participants draw on personally-experienced rather than received knowledge in performing everyday, real-world tasks.

4.4.3 The structural aspect

The structural aspect of validity refers to the level of fit between the internal structure of the test and the internal structure of the construct domain. It is related to the issue of construct representativeness we discussed earlier. A first step toward the goal of structural validity is taken by interviewing and eliciting knowledge from job incumbents in all areas that represent the performance domain. For example, in our study with military leaders, we interviewed officers in all three of the major branch categories within the Army (i.e., combat arms, combat support, combat service support). The goal of structural validity also is served by administering measurement instruments (e.g., the Tacit Knowledge Survey) to a wide variety of job incumbents. By using broad samples of job incumbents, we are able to avoid basing our analyses and test development on a restricted subset of the tacit-knowledge domain. Of course, the structural aspect of validity is addressed most directly through statistical techniques like cluster analysis and multidimensional scaling that identify the internal structure of the sample of items. By examining the internal structure we cast a wider net in our selection of tacit-knowledge items, and in so doing, we have improved our prospects for developing tacit-knowledge tests that mirror the structure of the construct domain (e.g., the domain of practical, action-oriented knowledge that individuals acquire from personal experience).

4.4.4 The generalizability aspect

The generalizability aspect of validity refers to the extent to which score properties and interpretations generalize across groups, settings, and tasks. The generalizability aspect includes concerns that traditionally fall under the heading of "reliability." In the context of tacit-knowledge test development, the goal of generalizability calls for tacit-knowledge test scores that generalize across (1) roles within the organization, (2) repeated administrations, and (3) alternate forms of the test. Test development efforts relevant to the content, substantive, and structural aspects of validity also are relevant to the generalizability aspect. In general, by seeking to specify and measure the construct, rather than merely pursuing correlation with an external criterion, we presumably increase the generalizability of score interpretations for our tacit-knowledge tests.

4.4.5 The external aspect

The external aspect of validity refers to the issue of criterion-related validity. That is, we seek to establish that the test relates to an external criterion. More specifically, the goal is to obtain evidence of convergent and discriminant validity. Establishing criterion-related validity entails showing that tacit-knowledge test scores correlate more highly (i.e., converge) with theoretically related constructs (e.g., performance) and correlate less highly (i.e., diverge) with theoretically distinct constructs (e.g., general cognition, formal job knowledge).

Test-development efforts to specify and measure the tacit-knowledge construct also support the goal of criterion validity. For example, job incumbents are asked to provide examples of important lessons they learned in the course of performing their job rather than knowledge they gained in school. These instructions increase the likelihood that the tacit-knowledge items obtained will be related to performance criteria and be distinct from formally-acquired knowledge. Research during the item-selection phase involves assessing more directly the relation of these items to external criteria. This step helps to identify tacit-knowledge items that are indicative of successful performance.

Beyond these efforts during test development, additional steps should be taken to provide evidence of convergent and discriminant validity. For tacit-knowledge tests, possible discriminant evidence would be that which discounts the effects of general cognition, reading comprehension, and formally-acquired knowledge on tacit-knowledge test scores. Evidence of convergent validity would include a correlation between tacit-knowledge test scores and variables such as perceived job effectiveness, degree and rate of career advancement, and performance on job-relevant tasks. To obtain such evidence requires conducting a validation study in which measures of these variables are administered to or obtained from individuals. For example, in our research with managers and military leaders, we administered the tacit-knowledge inventory along with a measure of general cognition and related constructs, and obtained various performance criteria, such as supervisor ratings, salary and productivity. Correlational and hierarchical regression analyses can be used to assess convergent and discriminant validity. Convergent validity is supported by a significant relationship between tacit-knowledge test scores and the performance criterion (e.g., supervisor ratings). Discriminant validity is supported by zero to moderate correlations with measures such as general cognition and general job knowledge, as well as the incremental validity of tacit-knowledge test scores beyond these measures.

4.4.6 The consequential aspect

The consequential aspect of validity refers to the value implications of the intended use of score interpretation as a basis for action. Because tacit-knowledge tests may be used for employee assessment and development, or even selection, it is important to consider how the knowledge included in those tests fits into the culture and rules of the organization. For example, if an item of tacit knowledge meets all the criteria discussed above (e.g., satisfies the definition of tacit, exhibits a strong positive correlation with effective performance), but it conflicts with the organizational culture (e.g., suggesting that females should be given less responsibility than males) or it involves disobeying a regulation (e.g., suggesting that financial figures should be fudged when information is unavailable), then it may be inappropriate to include the item in a tacit-knowledge test. Relying on experts to review the tacit-knowledge items throughout the test-development process helps to ensure that issues related to the consequential aspect of validity are addressed.

4.5 Summary

The goal of the test-development process outlined in this section is to support the construction of valid tacit-knowledge tests. Our theoretical model of tacit knowledge, described in the previous section, constitutes, we believe, a step in the direction of this goal. By elaborating on what we consider to be tacit knowledge at a theoretical level, we set the stage for a more detailed consideration of item content during the selection process and, in so doing, increase the substantive validity of our tests. The analysis of item ratings and performance data constitutes a second step towards measuring the construct. By identifying those items with the strongest association with performance criteria, we increase the probability that we will select items and construct test questions that embody the construct—given that tacit knowledge has clear benefits for performance. The analysis of the underlying structure by sorting items into categories constitutes a third step toward our goal. By examining the structure of the tacit-knowledge space (based on the data from our sample), we are able to make more informed decisions about the distribution of item content in our tacit-knowledge tests and, in so doing, increase the structural validity and generalizability of score interpretations. Finally, by conducting validation studies we provide support that tacit knowledge is relevant to understanding performance in the domain of interest and that it contributes to that understanding beyond traditional indicators of performance. In the next two sections we discuss the development and validation of tests to measure tacit-knowledge in civilian and military domains.

5. The role of practical cognition in everyday settings

Our program of research is based on the notion that there is more to successfully predicting performance than just measuring the so-called general factor from conventional psychometric tests of cognition (see Sternberg and Wagner, 1993). We propose that tacit knowledge, as an aspect of practical cognition, is a key ingredient of success in any domain. Of course, there are those who disagree with this position (see Jensen, 1993; Ree and Earles, 1993; Schmidt and Hunter, 1993, 1998), suggesting that individual differences in performance are explained primarily by general cognitive skill. Some proponents of using general cognitive skill tests argue further that the value of these tests are that they are applicable for all jobs, have lowest cost to develop and administer, and have the highest validity (e.g., Schmidt and Hunter, 1998). But even Schmidt and Hunter acknowledge that alternative measures such as work sample tests and job knowledge tests have comparable and perhaps even higher validities than general skill tests, and provide incremental prediction above the latter.

A program of research by Sternberg and his colleagues has conducted tacit-knowledge research with business managers, college professors, elementary-school students, sales people, college students, and general populations. This important aspect of practical cognition, in study after study, has been found generally to be uncorrelated with academic cognition as measured by conventional tests, in a variety of populations, occupations, and at a variety of age levels (Sternberg, et al., 1993; Sternberg et al., 1995; Wagner, 1987; Wagner and Sternberg, 1985). A major task of this tacit-knowledge research has been to identify the content of tacit knowledge and develop ways to measure the possession of tacit knowledge. Tacit-knowledge tests present a set of problem situations and ask respondents to rate the quality or appropriateness of a number of possible responses to those situations. (The format and development of tacit-knowledge tests were discussed in the previous section.) In this section, we review the tacit-knowledge studies that have been conducted in civilian settings and in the next section, we present a specific example of a tacit-knowledge project with military leaders.

5.1 Academic psychologists

One of the first studies in the program of tacit-knowledge research was conducted by Wagner and Sternberg (1985) with academic psychologists. Wagner and Sternberg developed a test of tacit knowledge for academic psychologists based on interviews with five full professors and administered the test to three groups. The first group consisted of 54 faculty members from 20 psychology departments, identified as either among the top fifteen nationally ranked colleges or outside the top fifteen. The second group consisted of 104 psychology graduate students from the same departments as the faculty members. The third group consisted of 29 Yale undergraduates. Each participant was given 12 work-related situations, each with from 6 to 20 response options. For example, one questions described a second-year assistant professor who in the past year had published two unrelated empirical articles, who had one graduate student working with him, and who had not yet received external funding. His goal was to become a top person in his field and get tenure in his department. Participants were asked to rate on a scale from 1 to 9 the value of several pieces of advice regarding what the professor could do in the next two months. Examples of advice include: (1) improve the quality of his teaching, (2) write a grant proposal, (3) begin a long-term research project that might lead to a major theoretical article, (4) concentrate on recruiting more students, (5) serve on a committee studying university-community relations, and (6) begin several related short-term projects, each of which may lead to an empirical article.

Responses to the test were scored by correlating ratings on each item with an index variable for group membership (1 = undergraduate, 2 = graduate student, 3 = faculty member). A positive correlation between item and group membership indicated that higher ratings on the item were associated with more expertise, and a negative correlation indicated the opposite. Wagner and Sternberg (1985) validated the test using several criteria. They obtained from faculty members citation rates, the number of publications, number of conferences attended in the last year, number of conference papers presented, distribution of time across teaching and research, academic rank, year Ph.D. was obtained, and level of institutional affiliation (high or low). For undergraduates, they obtained scores on the Verbal Reasoning section of the Differential Aptitude Test (Bennett, Seashore, and Wesman, 1974).

Wagner and Sternberg (1985) found that tacit-knowledge test scores correlated significantly, and positively, with number of publications (.33), number of conferences attended (.34), rated level of institution (.40), and proportion of time spent in research (.39). For the undergraduates, tacit-knowledge tests scores did not correlate significantly with verbal-reasoning scores ($r = -.04$, ns).

In a follow-up study by Wagner (1987), a revised version of the test was administered to 91 faculty, 61 graduate students, and 60 Yale undergraduates. The revised test contained 12 situations with 9 to 10 response options. Wagner obtained ratings for both conceptions of what the person would do in their actual job and what they would do in their ideal job. Scores were obtained for the overall test, and for six subscales that crossed three kinds of tacit knowledge: tacit knowledge about managing oneself, managing others, and managing tasks, with two orientations of tacit knowledge: local (pertaining to the situation at hand) versus global (pertaining to a bigger picture) tacit knowledge.

A different scoring method was used than in Wagner and Sternberg (1985). An expert profile was created by administering the test to a sample of professors who were nominated as high on practical cognition. A distance score (d^2) was computed between the participant's ratings and the mean of the experts' ratings. The mean d^2 values for the three groups were 339 for faculty, 412 for graduate students, and 429 for undergraduates, indicating that tacit knowledge increased, on average, with level of experience (a smaller value representing greater tacit knowledge). There were exceptions in each group, however, suggesting that what mattered was not merely experience by what one has learned from experience.

Wagner then examined the relationship of tacit knowledge with the same criterion measures that were used in Wagner and Sternberg (1985). Because the tacit-knowledge test was scored using a distance measure, a lower distance, or smaller value, represents better tacit-knowledge score. Therefore, negative correlations reflect a positive association between tacit-knowledge scores and the criterion.

For the actual-job ratings, significant correlations were obtained between tacit-knowledge scores and ratings of department (-.48), number of citations (-.44), number of publications (-.28), proportion of time spent on research (-.41), and number of papers presented. The correlations for ideal-job ratings were slightly lower, but comparable. Again, the tacit-knowledge scores did not correlate with verbal-reasoning skill. Wagner did find significant intercorrelations among the six subscales, ranging from .2 to .4. He interpreted these correlations to indicate a weak general factor for tacit knowledge, a factor that appears to be distinct from the general factor measured by traditional cognition tests.

5.2 Business managers

Wagner and Sternberg (1985) and Wagner (1987) conducted studies with business managers in parallel to the studies with academic psychologists described above. That is, they involved similar methods but with a different performance domain.

Wagner and Sternberg (1985) developed a tacit-knowledge test for business managers based on interviews with 5 experienced and successful mid-level managers. The test consisted of 12 work-related situations with 9 to 20 response options and was administered to 54 managers (19 of whom were from among top 20 Fortune 500 companies), 51 graduate students from 5 business schools varying in prestige, and 22 Yale undergraduates. The criteria obtained for the managers included status in or outside the top Fortune 500 companies, number of years of management experience, number of years of formal schooling, salary, number of employees supervised, and level of job title. Undergraduates completed the DAT Verbal Reasoning subtest.

Responses to the test were scored by correlating ratings on each item with an index variable for group membership (1=undergraduates, 2=business school graduate students, 3=business managers). Wagner and Sternberg found significant correlations between tacit-knowledge and company level (.34), number of years of schooling (.41), and salary (.46). For the undergraduates, the correlation between tacit-knowledge scores and verbal-reasoning skill was not significant (.16), and again indicated that the tacit knowledge test was not a proxy for a traditional general cognition test.

In the second study, Wagner (1987) administered the test to 64 business managers, 25 business graduate students, and 60 Yale undergraduates. The distance scoring method, described above, was used. An expert profile was created from the responses of 13 business executives from Fortune 500 firms. The mean tacit-knowledge scores were 244 for business managers, 340 for business graduate students, and 417 for undergraduates, indicating greater tacit knowledge with experience. Correlations with the criterion measures were lower than those for academic psychologists. However, a significant correlation was obtained between tacit-knowledge scores and the number of years of management experience (-.30). Other correlations were in the predicted direction, but not significant. There was no significant correlation between tacit-knowledge scores and verbal-reasoning scores. And again, the six subscales generally correlated significantly with one another, with values ranging from .2 to .5, indicating a weak general factor for tacit knowledge.

In this study, the undergraduate participants completed the tacit-knowledge tests for both academic psychologists and business managers. The correlation between scores on the two tests was .58 and highly significant. Wagner concluded that not only do the subscales of the tacit-knowledge test correlate within a domain, but tacit-knowledge also appear to correlate across domains.

5.3 Center for creative leadership study

Further research on what later was formalized as the Tacit Knowledge Inventory for Managers (TKIM; Wagner and Sternberg, 1991) was conducted with a sample of 45 business executives who were participants in a Leadership Development Program (LPD) at the Center for Creative Leadership (Wagner and Sternberg, 1990). The purpose of the study was to validate the test against a managerial simulation and assess its discriminant validity with a variety of psychological measures. Wagner and Sternberg (1990) administered the TKIM with 9 work-related scenarios, each with 10 response options. Participants also completed, as part of the program, the Shipley Institute for Living Scale, a cognition test; the California Psychological Inventory, a self-report

personality inventory; the Fundamental Interpersonal Relations Orientation-Behavior (FIRO-B), a measure of desired ways of relating to others; the Hidden Figures Test, a measure of field independence; the Myers-Briggs Type Indicator, a measure of cognitive style; the Kirton Adaptation Innovation Inventory, a measure of preference for innovation; and the Managerial Job Satisfaction Questionnaire. The participants' behavior was also assessed on two managerial simulations.

Beginning with zero-order correlations, the best predictors of managerial performance on the simulation were tacit knowledge (r = -.61, p < .001) and overall cognitive ability (r = .38, p < .001). (The negative correlation for tacit knowledge reflects the deviation scoring system used, in which better performance corresponds to less deviation from the expert prototype and thus to lower scores.) The correlation between tacit knowledge and overall cognitive ability was not significant (r = -.14, p > .05).

Hierarchical regression analyses were performed to examine the unique predictive value of tacit knowledge when used in conjunction with the various other cognition and personality tests. For each hierarchical regression analysis, the unique prediction of the TKIM was represented by the change in R^2 from a restricted model to a full model. In each case, the restricted model contained a subset of all the measures, and the full model was created by adding the TKIM to the equation. A significant change in R^2 indicated that the predictive relation between tacit knowledge and the simulation performance was not accounted for by the set of predictors in the restricted model.

In every case, tacit knowledge accounted for a significant increases in variance. In addition, when tacit knowledge, IQ, and selected subtests from the personality inventories were combined as predictors, nearly all of the reliable variance in the criterion was accounted for. These results support the strategy of enhancing validity and utility by supplementing existing selection procedures with additional ones. They also suggest that the construct of tacit knowledge cannot readily be subsumed by the existing constructs of cognitive skill and personality represented by the other measures used in the study.

5.4 Salespeople

Two studies were conducted by Wagner, Rashotte and Sternberg (1994; see also Wagner, Sujan, Sujan, Rashotte, and Sterberg, 1999) with salespeople. The objective of the first study was to develop and validate a "rules-of-thumb" approach to measuring tacit knowledge. Previous studies relied on empirical scoring, using either the correlation between items and an index of group membership or the deviation from an expert profile. Wagner et al. sought to identify a more objective, expert-based scoring method based on the rules of thumb that salespeople use to optimize their performance.

Based on a interviews, literature on sales, and personal experience, these investigators generated a list of rules of thumb for salespeople. The rules of thumb were divided into several categories, such as setting sales goals, handling the customer who stalls, attracting new accounts, and handling the competition. In the category of attracting new accounts, examples of rules of thumb included (1) be selective in regard to whom you direct your promotion efforts and (2) ask your customers to provide leads to new accounts.

The sample consisted of two groups. The first group consisted of salespeople with an average 14 years sales experience. The second group consisted of 50 undergraduates at Florida State University. The participants were administered eight sales scenarios, with 8 to 12 response options constructed by the rules-of-thumb approach. The options included accurate representations of the rules of thumb as well as weakened

or distorted versions of them. Responses were evaluated based on the extent to which participants preferred the actual or distorted versions of the rules of thumb. In addition to the sales test, the undergraduates completed the DAT Verbal Reasoning test.

Participants were asked to rate the appropriateness of each strategy for addressing the problem. Points were awarded based on the participant's endorsement of the actual rules of thumb. Wagner et al. found that scores on the tacit-knowledge test improved with experience. The average score for salespeople was 209 versus 166 for undergraduates. The total scores for undergraduates were uncorrelated with verbal-reasoning test scores.

In the second study, measures of sales performance were obtained in addition to tacit-knowledge test scores. Participants included 48 life-insurance salespeople with an average of 11 years sales experience and 50 undergraduates at Florida State University with no sales experience. Participants in both groups completed the TKIS, and undergraduates completed the DAT Verbal Reasoning test. In addition, the investigators obtained from the salespeople data on the number of years with the company, number of years in sales, number of yearly quality awards, yearly sales volumes and premiums, college background, and business education.

Tacit knowledge again increased with experience, with the scores 165 and 206 for undergraduates and salespeople respectively. Significant correlations were obtained between tacit-knowledge scores and number of years with the company (.37), number of years in sales (.31), number of yearly quality awards (.35), and business education (.41). When local and global scores were also computed, Wagner et al. found that global tacit-knowledge scores also correlated significantly with yearly sales volumes and premiums (*r*s ranging from .26 to .37). The tacit-knowledge scores again did not correlate significantly with verbal-reasoning scores.

5.5 Air force recruits

In a study carried out at the Human Resources Laboratory at Brooks Air Force Base under the supervision of Malcolm Ree, Eddy (1988) examined relations between the TKIM and the Armed Services Vocational Aptitude Battery (ASVAB) for a sample of 631 Air Force Recruits, 29 percent of whom were females, and 19 percent of whom were members of a minority group. The ASVAB is a multiple-aptitude battery used for selection of candidates into all branches of the United States Armed Forces. Prior studies of the ASVAB suggest that it is a typical measure of cognitive skill, with correlations between ASVAB scores and other cognitive skill measures of about .7. Factor-analytic studies of the ASVAB also suggest that it appears to measure the same verbal, quantitative, and mechanical skills as the Differential Aptitude Tests, and the same verbal and mathematical knowledge as the California Achievement Tests.

Eddy's (1988) study showed small correlations between tacit knowledge and ASVAB subtests. The median correlation was -.07, with a range from .06 to -.15. Of the 10 correlations, only two were significantly different from zero, despite the large sample size of 631 recruits. A factor analysis of all the test data, followed by oblique rotations, yielded the usual four ASVAB factors (vocational-technical information, clerical/speed, verbal skill, and mathematics) and a distinct tacit-knowledge factor. The factor loading for the TKIM score on the tacit-knowledge factor was .99, with a maximum loading for scores on the four ASVAB factors of only .06. Upon oblique rotation, the four ASVAB factors were moderately intercorrelated, but the correlations between the tacit knowledge factor and the four ASVAB factors were near zero (.075, .003, .096, .082).

An additional point about these results concerns the possibility that measures of tacit knowledge might identify potential managers from nontraditional and minority backgrounds whose practical knowledge suggests that they would be effective managers, even though their performance on traditional selection measures such as cognition tests does not. Eddy (1988) did not report scores separately by race and sex, but did report correlations between scores and dummy variables indicating race and sex. Significant correlations in the .2 to .4 range between ASVAB subtest scores and both race and sex indicate that on the ASVAB, minority-group members scored more poorly than majority group members, and women scored more poorly than men. Nonsignificant correlations between tacit knowledge and both race (.03) and sex (.02), however, indicate comparable levels of performance on the tacit-knowledge measures between minority and majority-group members and between females and males.

5.6 Managers across organizational levels

In a study focusing on the development of tacit knowledge over the managerial career, Williams and Sternberg (cited in Sternberg et al., 1995) constructed a measure of both a general and a level-specific tacit-knowledge. They obtained nominations from superiors for "outstanding" and "underperforming" managers at the lower, middle, and upper levels in four high-technology manufacturing companies. This approach allowed them to delineate the specific content of tacit knowledge for each level of management (lower, middle, and upper) by examining what experts at each level knew that their poorly-performing colleagues did not.

Williams and Sternberg identified specialized tacit knowledge for each of the three management levels and found that this knowledge was differentially related to success. These results were derived from comparing responses of outstanding and underperforming managers within each management level on level-specific tacit-knowledge inventories. Within the domain of intrapersonal tacit knowledge, knowledge about how to seek out, create, and enjoy challenges is substantially more important to upper-level executives than to middle- or lower-level executives. Knowledge about maintaining appropriate levels of control becomes progressively more significant at higher levels of management. Knowledge about self-motivation, self-direction, self-awareness, and personal organization is roughly comparable in importance at the lower and middle levels, and somewhat more important at the upper level. Finally, knowledge about completing tasks and working effectively within the business environment is substantially more important for upper-level managers than for middle-level managers, and substantially more important for middle-level managers than for lower-level managers. Within the domain of interpersonal tacit knowledge, knowledge about influencing and controlling others is essential for all managers, but especially for those at the upper level. Knowledge about supporting, cooperating with, and understanding others is extremely important for upper-level executives, very important for middle-level executives, and somewhat important for lower-level executives.

In addition, Williams and Sternberg examined the relationship of tacit knowledge with several criteria across levels. They found that tacit knowledge was related to the following measures of managerial success: compensation ($r = .39$, $p < .001$), age-controlled compensation ($r = .38$, $p < .001$), and level of position ($r = .36$, $p < .001$). These correlations were computed after controlling for background and educational experience. Tacit knowledge was also moderately associated with enhanced job satisfaction ($r = .23$, $p < .05$).

These investigators further found that age, years of management experience, and years in current position were unrelated to tacit knowledge. The lack of a correlation of tacit knowledge with years of management experience suggests that it is not simply experience that matters, but perhaps what a manager learns from experience. A manager's years with current company was negatively related to tacit knowledge (r = -.29, p < .01), perhaps indicating that ineffective managers stayed around longer than effective managers. The number of companies that a manager had worked for was positively correlated with tacit-knowledge scores (r = .35, p < .001). Years of higher education was highly related to tacit knowledge (r = .37, p < .001), as was self-reported school performance (r = .26, p < .01). Similarly, college quality was related to tacit knowledge (r = .34, p < .01). These results in conjunction with the independence of tacit knowledge and overall cognitive ability suggest that tacit knowledge overlaps with the portion of these measures that are not predicted by overall cognitive ability.

Williams and Sternberg also performed hierarchical regression analyses to examine whether tacit knowledge contained independent information related to success that was distinct from that provided by background and experience. The pattern of results was similar across analyses. In the regression analysis predicting maximum compensation, the first variable entered in the regression equation was years of education, accounting for 19% of the variance (p < .001). The second variable entered was years of management experience, accounting for an additional 13% of the variance (p < .001). The third and final variable entered was tacit knowledge, accounting for an additional 4% of the variance (p = .04), and raised the total explained variance to 36%. In the regression predicting maximum compensation controlling for age, the number of years of education was entered into the equation first, accounting for 27% of the variance (p < .001). And second, tacit knowledge was entered, explaining an additional 5% of the variance (p = .03). This final regression demonstrates the value of tacit knowledge to managers who are relatively successful for their age.

5.7 College students

Williams and Sternberg (cited in Sternberg et al., 1993) studied the tacit knowledge of college students. They asked 50 Yale undergraduates the question: "What does it take to succeed at Yale that you don't learn from textbooks?" and used the responses to develop a tacit-knowledge inventory for college students. The inventory consisted of 14 situations and asked respondents to rate the quality of several options on a 1 to 9 scale. For example, one question described a student enrolled in a large introductory lecture course. The class requirements included three exams and a final. Participants were asked to rate how characteristic of their behavior it was to spend time doing various activities, such as (1) attending class regularly, (2) attending optional weekly review sections with a teaching fellow, (3) reading assigned text chapters thoroughly, (4) taking comprehensive class notes, and (5) speaking with the professor after class and during office hours.

The criteria were two indices: an academic index and an adjustment index. The academic index was a composite of high school GPA, college GPA, SAT scores, and CEEB achievement test scores. The adjustment index was a composite of a measure of happiness in college, a measure of self-perceived success in college, a measure of self-perceived success in using tacit knowledge, a measure of the extent of benefit each participant had experienced from acquiring tacit knowledge, and a measure of the rated closeness of the college to the participant's ideal college.

The academic and adjustment indices were not significantly correlated (-.09). Individual items of tacit knowledge correlated differently with the academic index and the adjustment index. The academic index was correlated with the perceived importance of maintaining a high GPA (.42); doing extra reading and school work not specifically assigned (.27); not attending optional weekly review sections (.23); not skimming required reading the morning before class (.37); not preparing a brief outline of points to raise in class discussion (.31); not helping friends with their assignments (.34); not behaving consistently from situation to situation (.25); finding it uncharacteristic to accept pressure and stress as parts of life (.30); finding it uncharacteristic to stand up for oneself (.34); and finding it uncharacteristic to play a sport or exercise regularly (.45).

Items that correlated significantly with the adjustment index included beliefs that professors value a clear, direct writing style, good organization of thoughts and ideas, and creative or unusual ideas (.38); beliefs that professors value papers that bring in outside interests or material (.27); beliefs that it is important sometimes to take on too many responsibilities at once (.31); seeking advice from several faculty in addition to one's own professors (.31); taking classes that permit occasional absences (.36); being positive and looking on the bright side of life (.42); not being intimidated (.33); being flexible (.27); maintaining a strong sense of confidence and independence (.37); not worrying unnecessarily or destructively (.31); knowing how to make oneself happy (.32); and not letting small disappointments affect one's long-term goals (.29).

Williams and Sternberg also obtained prediction of academic and adjustment indices with subsets of items from the tacit-knowledge inventory. Four items (not preparing an outline of points to raise in class discussion; maintaining a high GPA; not helping friends with assignments; and not playing a varsity or intramural sport) were predictive of the academic index, with an overall R^2 of .43. Six items (believing professors value a clear, direct writing style; maintaining a strong sense of confidence and independence; standing up for oneself; sometimes taking on too many responsibilities at once; seeking advice from faculty in addition to the course instructor; and taking classes that permit occasional absences) were predictive of the adjustment index, with an overall R^2 of .63. This study showed that tacit knowledge is important not only in occupational settings, but in school settings as well.

5.8 Conclusions from the tacit-knowledge research program

We organize a discussion of the findings from the tacit-knowledge research around four main issues: (a) the relationship of tacit knowledge to experience; (b) the relationship of tacit knowledge to general cognition; (c) tacit knowledge as a general construct; and (d) the relationship of tacit knowledge to performance.

5.8.1 Tacit knowledge and experience

In most of the studies reviewed above, tacit knowledge was found to relate to experience, indicated either by group membership (expert versus novice), or the number of years in one's current position.

In several studies, Sternberg and his colleagues showed that individuals with less experience in a given domain exhibit lower tacit-knowledge scores (Wagner, 1987; Wagner and Sternberg, 1985; Sternberg et al., 1993). In Wagner and Sternberg (1985), for example, group differences were obtained between business managers, business graduate students, and undergraduates on 39 of the response-item ratings on a tacit-knowledge test for managers, with a binomial test of the probability of finding this

many significant differences by chance yielding $p < .001$. Comparable results were obtained with Yale undergraduates, psychology graduate students, and psychology faculty on a tacit-knowledge test for academic psychologists. In addition, Wagner (1987) found that business managers obtained the highest tacit knowledge scores followed by business graduate students, and undergraduates, with comparable results obtained in a study with psychology professors, psychology graduate students, and undergraduates. Wagner et al. (1994) also found that scores on a tacit-knowledge test for salespeople correlated significantly with number of years of sales experience.

Williams and Sternberg (cited in Sternberg et al., 1995), however, did not find significant correlations between several experience-based measures, including age, years of management experience, and years in current position, and tacit-knowledge scores. But they did find that the importance of specific pieces of tacit knowledge varied across organizational level. Their findings suggest that it may not simply be the amount of experience but what a manager learns from experience that matters to success.

5.8.2 Tacit knowledge and general cognition

In proposing a new approach to measuring cognition, it is important to show that one has not accidentally reinvented the concept of "*g*," or so-called general skill, as measured by traditional cognition tests. We do not dispute the relevance of general cognitive skill to performance. Schmidt and Hunter (1998) have shown that *g* predicts performance in a number of domains. Our aim is to show that tacit-knowledge tests measure something in addition to *g*. In all the above studies in which participants were given a traditional measure of cognitive skill, tacit-knowledge test scores correlated insignificantly with *g*.

The most consistently used measure of *g* in the above studies was the Verbal Reasoning subtest of the DAT. The absolute values of the correlations between tacit knowledge and verbal reasoning ranged from .04 and .16 with undergraduate samples (Wagner, 1987; Wagner and Sternberg, 1985) and .14 with a sample of business executives (Wagner and Sternberg, 1990).

One potential limitation of these findings is that they were obtained with restricted samples (e.g., Yale undergraduates, business managers). However, similar support for the relationship between tacit knowledge and *g* was found in a more general sample of Air Force recruits studied by Eddy (1988). The correlations between scores on the TKIM and ASVAB scales were modest, and none of the four ASVAB factors correlated significantly with the tacit-knowledge factor.

Tacit-knowledge tests may also be a better predictor than measures of personality, cognitive style, and interpersonal orientation as suggested by the findings from the Center for Creative Leadership study (Wagner and Sternberg, 1990). Sternberg and Grigorenko recently developed a test of common sense for the work place (e.g., how to handle oneself in a job interview) that predicts self-ratings of common sense but not self-ratings of various kinds of academic skills. The test also predicts supervisory ratings at a correlational level of about 4.

Finally, there is evidence that tacit knowledge may even correlate negatively with measures of academic cognition and achievement in some environments. In a study in a rural village in Kenya, Sternberg et al. (in press) developed a test to measure children's tacit knowledge for herbal medicines used to treat various illnesses. Parasitic infections are endemic among this population, and knowledge of these medicines and how to use them is important to adaptation to the environment. This knowledge, however, is not acquired in the classroom, but rather in the community from family members and healers.

The tacit-knowledge test for herbal medicines consisted of brief stories describing the specific manifestations of a given illness and provided the child with options regarding how to treat the illness (see Sternberg et al., in press). The tacit-knowledge test, along with the Raven Colored Progressive Matrices test (Raven, 1958), the English Mill Hill Vocabulary Scale (Raven, Court, and Raven, 1992), Dholuo (home language) Vocabulary Scale, and school-based measures of English and math achievement, were administered to 85 children ages 12 to 15. The tests of academic cognition were all significantly and positively correlated with each other. Scores on the tacit-knowledge test correlated in a negative direction with all of the academic cognition tests, and showed a significant negative correlation with scores on the vocabulary tests. Tacit-knowledge scores also exhibited a significant negative correlation with English achievement. Sternberg et al. concluded that practical cognition, as manifested in tacit knowledge relevant to adaptation in daily life, may be distinct from the kind of academic cognition associated with school success. The negative correlation between tacit-knowledge scores and some of the academic-cognition measures supports the claim that expertise developed in one environment (e.g., school) may have limited application in other environments (e.g., home or community life). Thus, there is a growing body of evidence, in work, school and community settings, which suggests that tacit knowledge measures a distinct construct from general, academic cognition.

5.8.3 Tacit knowledge as a general construct

Although the kinds of informal procedural knowledge measured by tacit-knowledge tests do not correlate with traditional psychometric cognition, tacit-knowledge test scores do correlate across domains. Furthermore, the structure of tacit knowledge appears to be represented best by a single, general factor.

Wagner (1987) examined the structure of tacit knowledge inventory for managers. He performed two kinds of factor analyses on the tacit-knowledge scores of these business managers in his study. First, a principal-components analysis yielded a first principal component that accounted for 44 percent of the total variance, and 76 percent of total variance after the correlations among scores were disattenuated for unreliability. The 40 percent variance accounted for by the first principal component is typical of analyses carried out on traditional cognitive-skill subtests. Second, results of a confirmatory factor analysis suggested that a model consisting of a single general factor provided the best fit to the data. The results of both factor analyses suggested a general factor of tacit knowledge.

Similar analyses were performed on a measure of tacit knowledge for academic psychologists. Consistent with the manager study, the factor analytic results suggested a single factor of tacit knowledge within the domain of academic psychology. Wagner (1987) also examined the generalizability of tacit knowledge across domains by administering both tacit-knowledge measures (for business managers and academic psychologists) to undergraduates in his study. He obtained a significant correlation of .58 between the two scores, suggesting that in addition the existence of a general factor of tacit knowledge within a domain, individual differences in tacit knowledge generalize across domains. These findings lend support for a common factor underlying tacit knowledge; a factor that is considered to be an aspect of practical cognition

5.8.4 Tacit knowledge and performance

Finally, we have shown that tacitknowledge measures are predictive of performance in a number of domains, correlating between .2 to .5 with measures such as rated prestige of business or institution, salary, simulation performance, and number of publications.

These correlations, uncorrected for attenuation or restriction of range, compare favorably with those obtained for overall cognitive ability within the range of skills we have tested.

In studies with business managers, tacit-knowledge scores correlated in the range of .2 to .4 with criteria such as salary, years of management experience, and whether or not the manager worked for a company at the top of the Fortune 500 list (Wagner, 1987; Wagner and Sternberg, 1985). Wagner and Sternberg (1990) obtained a correlation of .61 between tacit knowledge and performance on a managerial simulation, and found that tacit-knowledge scores explained additional variance beyond overall cognitive ability and other personality and skill measures. In a study with bank branch managers Wagner and Sternberg (1985) obtained significant correlations between tacit-knowledge scores and average percentage of merit-based salary increase ($r = .48$, $p < .05$) and average performance rating for the category of generating new business for the bank ($r = .56$, $p < .05$).

Williams and Sternberg (cited in Sternberg et al., 1995) also found that tacit knowledge was related to several indicators of managerial success, including compensation, age-controlled compensation, level of position, and job satisfaction, with correlations ranging from .23 to .39.

Although much of the tacit-knowledge research has involved business mangers, there is evidence that tacit knowledge explains performance in other domains. In the field of academic psychology, correlations in the .4 to .5 range were found between tacit-knowledge scores and criterion measures such as citation rate, number of publications, and quality of department (Wagner, 1987; Wagner and Sternberg, 1985). In studies with salespeople, Wagner et al. (1994) found correlations in the .3 to .4 range between tacit knowledge and criteria such as sales volume and sales awards received. Finally, tacit knowledge for college students was found to correlate with indices of academic performance and adjustment to college (Williams and Sternberg, cited in Sternberg et al., 1993).

In summary, the program of tacit-knowledge research reviewed above shows that generally tacit knowledge increases with experience, but is not simply a proxy for experience; that tacit-knowledge tests measure a distinct construct from that measured by traditional, abstract cognition tests; that scores on tacit-knowledge tests represent a general factor, which appears to correlate across domains; and finally, that tacit knowledge tests are predictive of performance in a number of domains, and compare favorably with those obtained for overall cognitive ability within the range of skills we have tested.

6. An example of the application of the framework: The ALL practical-cognition study

In their everyday lives, people continually need to make situational judgments: how to get along with a difficult boss, how to break bad news to a friend or coworker, how to handle anger or disappointment after a failed endeavor. These skills are important to life adjustment in general, and to workplace efficacy, in particular. This report describes a project aimed at measuring such skills.

The project described here had three major goals. The first goal was to develop a theory-based instrument to measure practical cognition as measured by everyday situational-judgment skills. The second goal was to evaluate the psychometric properties of the instrument, including item characteristics, reliability, and both internal and external validity. The third goal was to compare the psychometric properties and the utility of the instrument in two cultural settings: the United States and Spain.

We also decided that there were certain goals that we explicitly were *not* setting for this project. The first was to measure all possible kinds of situational-judgment skills. In this first phase of our project, we concentrated on workplace situational-judgment skills. In later phases, we will seek to measure other kinds of situational-judgment skillsas well. The second thing we did not try to do was to measure everyday situational-judgment skills in an occupation-specific way. Previously, as described earlier, we had devised a number of inventories for specific occupations, such as managers, salespeople, university professors, university students, military officers, and elementary-school teachers (see Sternberg, Wagner, and Okagaki, 1993; Sternberg, Forsythe, et al., in press; Sternberg, Wagner, Williams, and Horvath, 1995; Wagner and Sternberg, 1986). We intended in this project to extend our methodology to jobs in general rather than devising yet another measure for another specific occupation. Third, we did not seek an inventory with "objectively correct" answers, because situational judgments are, by their nature, more or less useful or possibly justifiable, but they are not, strictly speaking, objectively correct or incorrect. The theoretical basis for our work is the triarchic theory of cognitive skills (Sternberg, 1985a, 1988, 1997, in press-b).

In the current work, we have sought to extend our past work in three ways. First, we have measured informal knowledge that is relatively more domain general than in our past work, where we have targeted specific jobs or career paths. Second, we have sought to extend our findings cross-culturally, using the same inventory in translated form in Spain as in the U.S. Third, we have used item-response-theory (IRT) scaling in order to explore the scalar properties of our inventory.

We also have used a new conceptual framework in this research, which is shown in Figure 1. This framework, used for item construction, crosses three objects of attention (dealing with self, dealing with others, dealing with tasks) with five categories of behavior (motivating, interpreting situations, behaving, following directions, and organizing).

Figure 1
Conceptual framework

	Motivating	Interpreting situations	Behaving	Following directions	Organizing
DS (self)					
DO (others)					
DT (tasks)					

6.1 Method

6.1.1 Materials

Two main kinds of materials were used in this project.

The Everyday Situational Judgment Inventory. The Everyday Situational-Judgment Inventory (ESJI) consists of descriptions of various situations encountered by many people. After each situation, there are 8 options for handling the situation. For each option listed, participants were asked to rate the quality of the option on a 1 (low) to 9 (high) Likert scale, where the anchor points were 1=extremely bad, 3=somewhat bad, 5=neither bad nor good, 7=somewhat good, and 9=extremely good. Participants were asked to select the number corresponding to their judgment, and to write it in the blank preceding each option. Participants were told that there was no one "right" answer—that the options were simply things that people might do in the situations described.

An example of an item is as follows:

You've been assigned to work on a project for a day with a fellow employee whom you really dislike. He is rude, lazy, and rarely does a proper job. What would be the best thing for you to do?

_____*Tell the worker that you think he is worthless.*

_____*Warn the worker that, if he is not "on his toes" today, you will complain to the supervisor.*

_____*Avoid all conversation and eye contact with the other worker.*

_____*Be polite to the other worker and try to maintain as business-like a manner as possible so that hopefully he will follow your example for the day.*

_____*Tell your supervisor that you refuse to work with this man.*

_____*The project is going to be impossible to accomplish with this worker, so you may as well not even try —you can always blame your bad work partner.*

_____*See if you can convince one of your friends to take your place and work with this employee.*

_____*Demand a raise from your supervisor; you should not have to tolerate these conditions.*

Participants were given as much time as they needed to finish the inventory.

Scoring for the ESJI was done in three ways:

1. *Profile matching (d^2).* For each of the 30 problems, a given respondent's responses were compared to the averaged (prototypical) responses to that problem. The following specific scoring procedure was used. For a given option, the difference between the participant's response and the sample-mean response was computed and squared. Squared differences were summed across the 8 options and averaged. Then the square root of this average was computed. The same procedure was repeated for each of the 30 items. Total score was the sum of these values.
2. *Rank-order correlation between individuals and mean profile (ρ).* For this measure, the rank orders of the responses for the mean profile of responses to a given item were correlated with the rank orders of the responses for an individual's profile of responses for that item. Thus, 8 observations for an individual were correlated with 8 observations for the mean profile. This analysis yielded a rank-order correlation (rho, or ρ) for each item. These correlations were averaged across the 30 problems.

3. *Dichotomized responses based on significance of* ***ρ.*** Dichotomized scores were created for each of the 30 items by assigning the item a score of 1 if the ρ value for that item was statistically significant and a 0 otherwise. In other words, the respondent got credit (1) if the respondent's item response pattern rank-order correlated significantly with the averaged item response pattern.

Performance evaluations. The performance of the individuals who filled out the ESJI was evaluated in two ways, via self-ratings and via supervisor ratings. In Part 1 of each evaluation (self and supervisor), ratings were coded in the following way: 1=definitely no, 5=not sure, 9=definitely yes. In Part 2, ratings were coded in the following way: 1=extremely bad, 3=somewhat bad, 5=neither bad nor good, 7=somewhat good, 9=extremely good.

Here are the items for self-ratings:

Part 1.

1. *My relationship with my supervisor is good.*
2. *My supervisor thinks highly of me.*
3. *I am satisfied with the development of my career.*
4. *I am planning a career change.*
5. *My relationships with my coworkers are good.*
6. *My greatest strength on the job is my ability to work well with others.*
7. *My greatest strength on the job is my ability to work independently.*
8. *My greatest strength on the job is my ability to manage tasks.*
9. *My greatest strength on the job is to motivate myself.*

Part 2.

1. *How would you rate your common-sense ability?*
2. *How would you rate your academic ability?*
3. *How would you rate your creative ability?*
4. *How good are you at working by yourself?*
5. *How good are you at working with others?*

The supervisors' evaluation rating scales were comparable.

Part 1.

1. *My relationship with this employee is good.*
2. *I think highly of this employee.*
3. *I am satisfied with this employee.*
4. *The employee's relationships with other coworkers are good.*

Part 2.

1. *How would you rate this employee's common sense ability?*
2. *How would you rate this employee's academic ability?*
3. *How would rate this employee's creative ability?*
4. *How would you rate this employee at working by him/herself?*
5. *How would you rate this employee at working with others?*
6. *How good is this employee at motivating him/herself?*
7. *How good is this employee at managing tasks?*
8. *How responsible is this employee?*

6.1.2 Participants

There were two sets of participants, from the United States and from Spain.

U.S. participants. There were 230 U.S. participants, 78 male, 149 female, 3 unidentified. The mean age was 35.8 years with a standard deviation of 13.5 years. The range was from 17 to 72 years of age. Mean time in the workplace was 6.7 years with a standard deviation of 7.9 years. Mean time in current position was 1.3 years with a standard deviation of 1.0 years. Job classifications of these participants included custodians, food-service workers in a dining hall, restaurant waiters and waitresses, salespeople, postal-service workers, taxi drivers, office personnel, and teachers.

Spanish participants. There were 227 Spanish participants, 112 male, 112 female, and 3 unidentified. The mean age was 36.1 years with a standard deviation of 9.8 years. The range was from 21 to 64 years of age. Mean time in the workplace was 7.6 years with a standard deviation of 8.6 years. Mean time in current position was 4.0 years with a standard deviation of 3.6 years. Job classifications of these participants clerks, bank office personnel, photography studio personnel, biology laboratory personnel, film developing studio personnel, lawyers' office support personnel, librarians, educational researchers, textbook editors, university teachers, air traffic controllers, administrative personnel of diverse institutions, and psychiatrists.

6.1.3 Design

The main dependent variables were responses to the performance-evaluation items (as answered by both workers and their supervisors). The main independent variables were scores on the ESJI.

6.1.4 Procedure

The ESJI administered individually in both the United States and Spain. The administration typically took 30-40 minutes, although the inventory was untimed. The instrument was constructed in the United States and then translated into Castilian Spanish by the Spanish team and checked for accuracy by a Spanish-speaking member of the U.S. team.

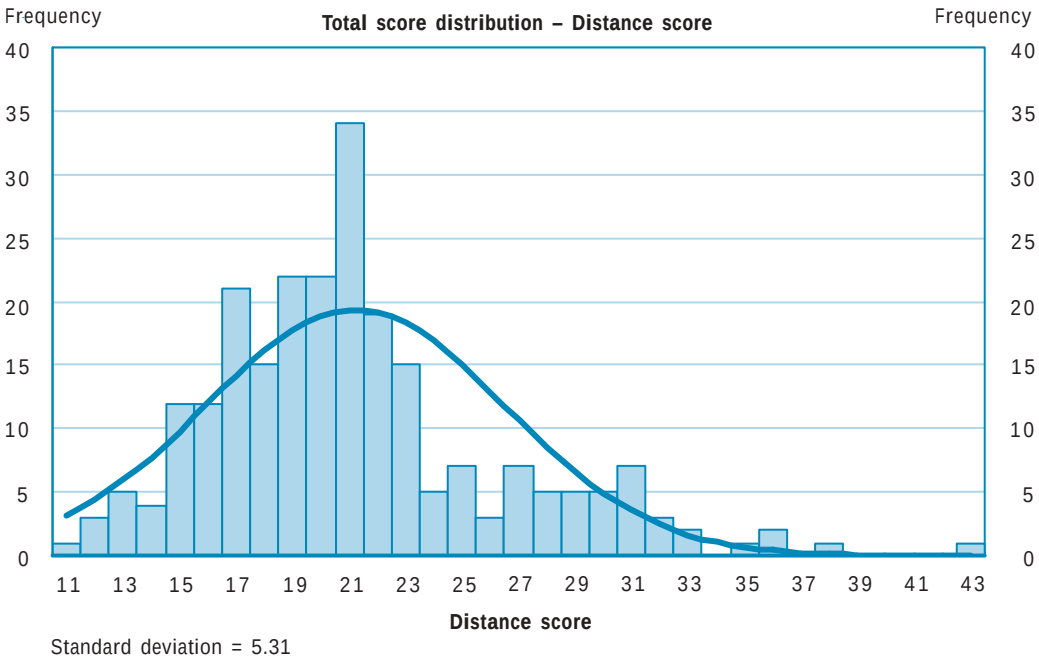
6.2 Results and discussion

6.2.1 Basic statistics and score distributions

Figure 2 (Panels A-F) shows the total-score distributions for each of the three methods of scoring for each of the samples (United States and Spain). Each figure also shows the mean and standard deviation for the given scoring system for the given sample, as well as the *N* on which the set of sample statistics is based.

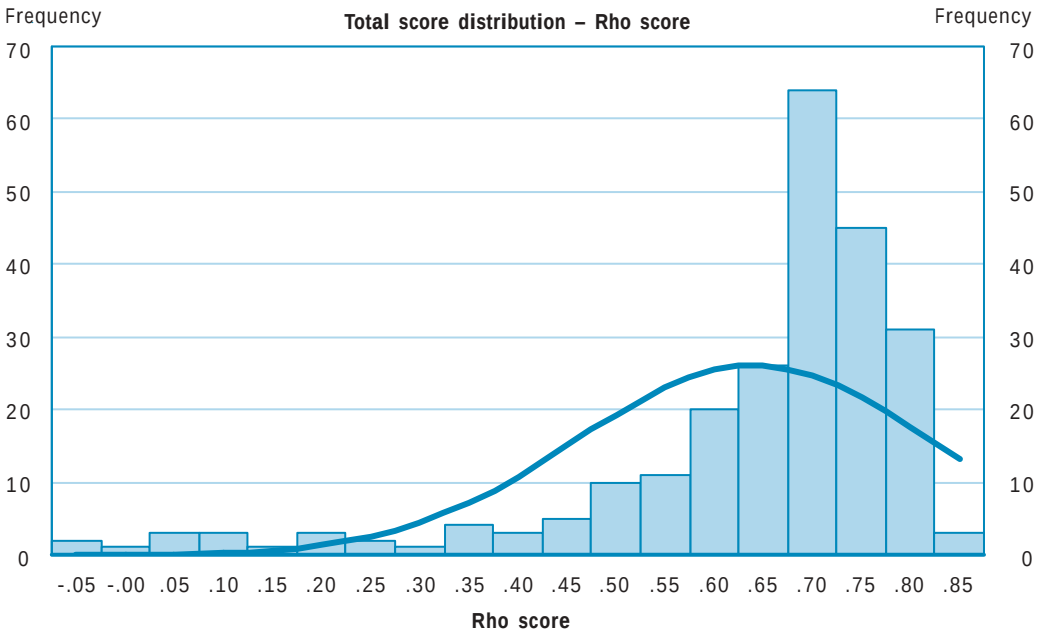
The Spanish sample performed better than did the U.S. sample, although because the occupations of the Spanish and U.S. samples were not exactly equated, the samples may not have been strictly comparable. The differences were statistically significant at the .001 level, regardless of the measure used. For the distance scores, t_{455} was equal to 8.47. For the rank-order correlation scores, the comparable *t*-value was 5.32, and for the 1/0 scores, 5.92.

Figure 2 (Panel A)
Total score characteristics: US sample



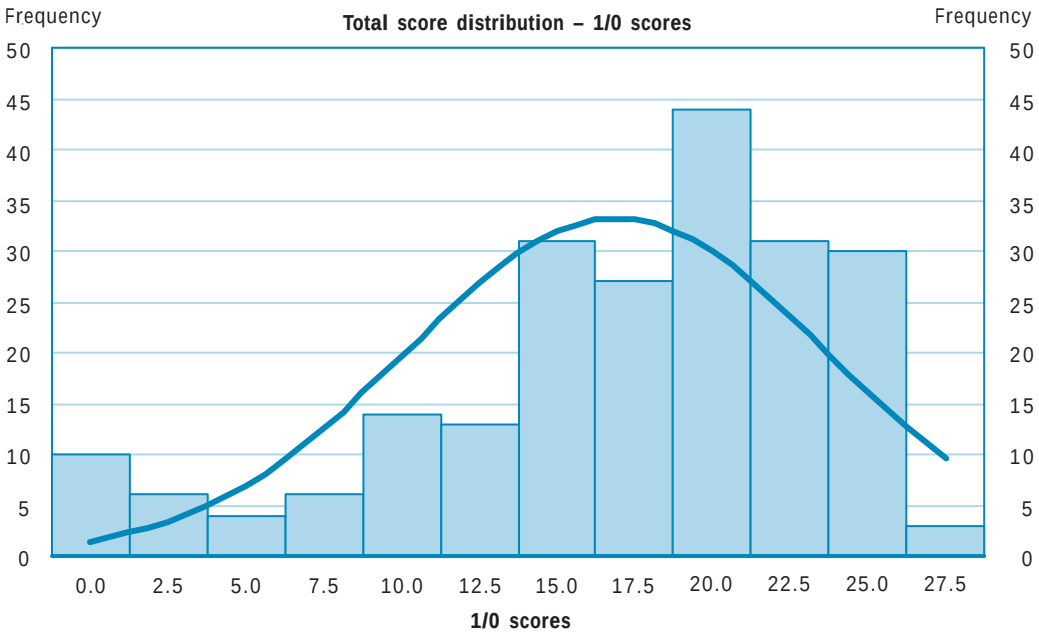
Standard deviation = 5.31
Mean = 21.2
Number = 230.00

Figure 2 (Panel B)
Total score characteristics: US sample



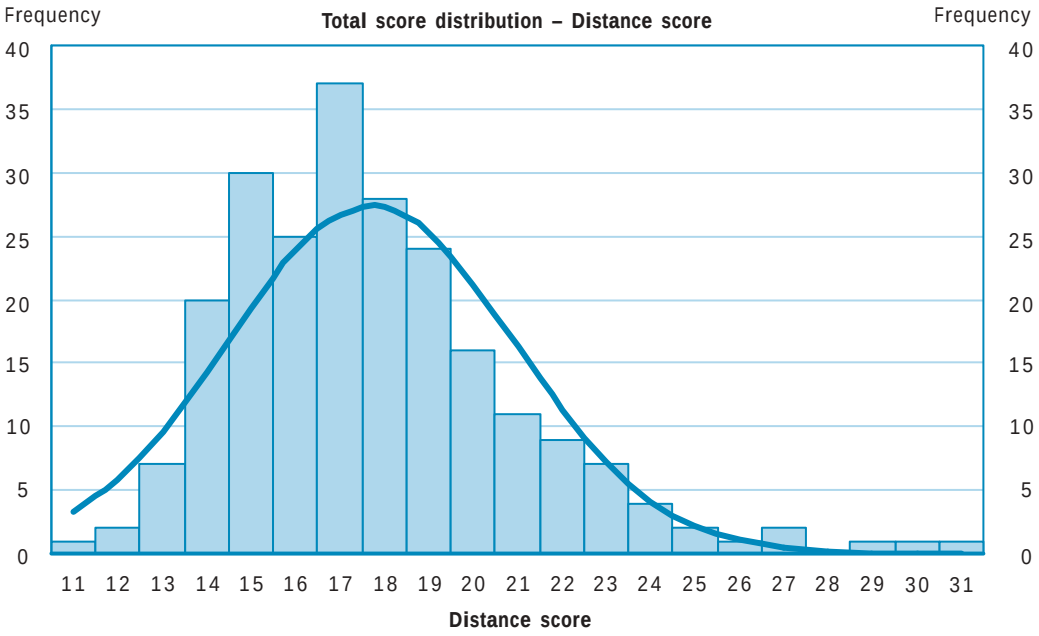
Standard deviation = .18
Mean = .64
Number = 230.00

Figure 2 (Panel C)
Total score characteristics: US sample



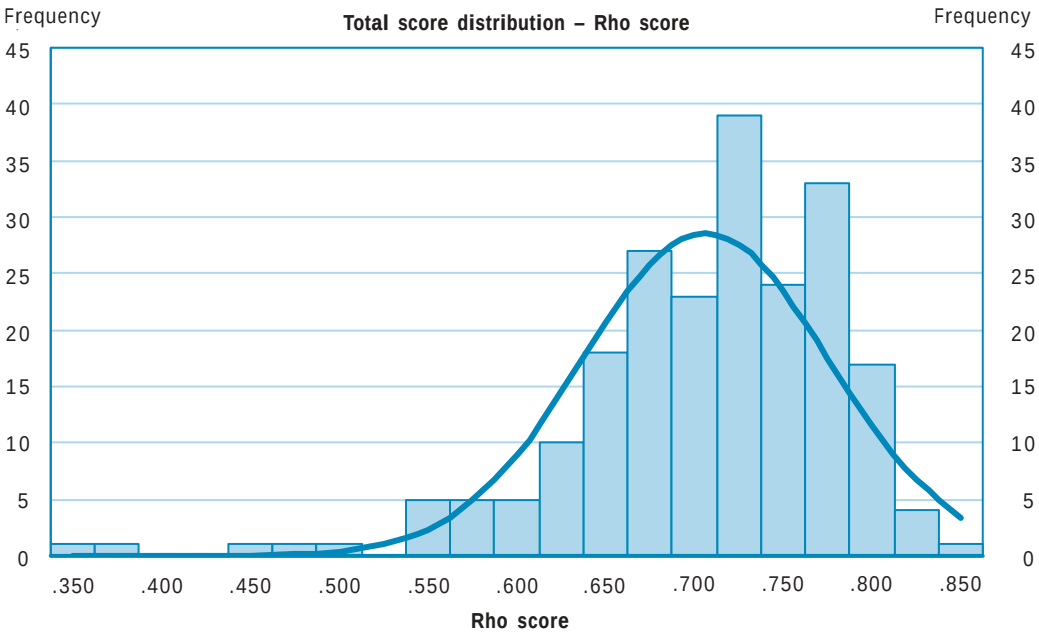
Standard deviation = 6.70
Mean = 16.9
Number = 230.00

Figure 2 (Panel D)
Total score characteristics: Spanish sample



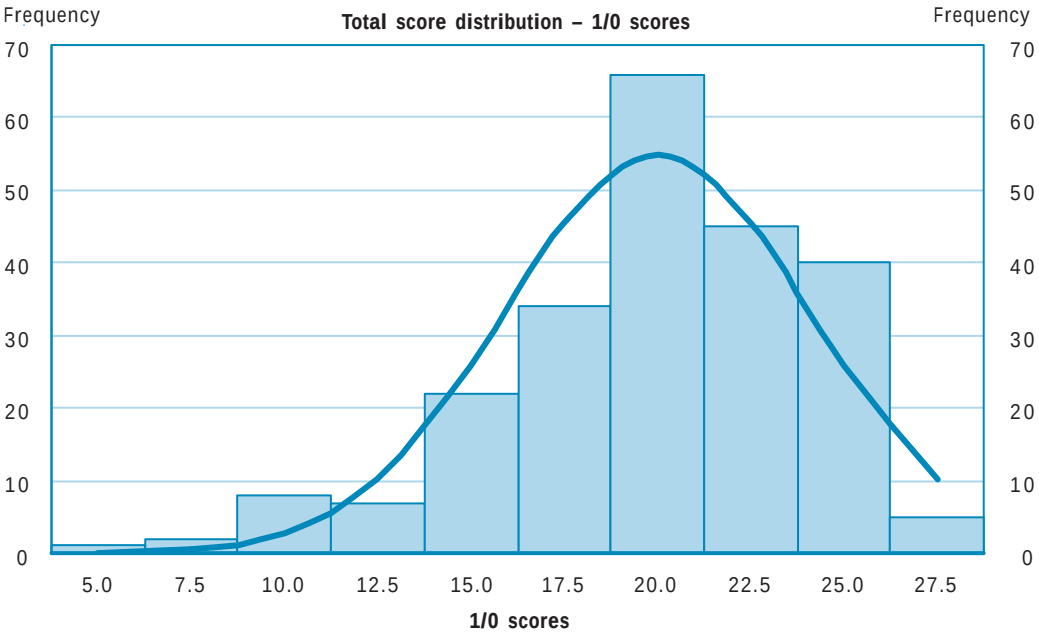
Standard deviation = 3.24
Mean = 17.7
Number = 227.00

Figure 2 (Panel E)
Total score characteristics: Spanish sample



Standard deviation = .07
Mean = .706
Number = 227.00

Figure 2 (Panel F)
Total score characteristics: Spanish sample



Standard deviation = 4.07
Mean = 20.0
Number = 227.00

6.3 Internal validation

6.3.1 Internal-consistency reliabilities

For distance scores, coefficient a internal-consistency reliabilities were .96 for the U.S. sample and .92 for the Spanish sample. For rank-order correlation scores, internal-consistency reliabilities were .94 for the U.S. sample and .73 for the Spanish sample. For 1/0 scores, KR-20 internal-consistency reliabilities were .89 for the U.S. sample and .62 for the Spanish sample. The overall 1/0 internal-consistency reliability for the combined samples was .82.

6.3.2 Comparison of responses between U.S. and Spanish samples

The correlation for mean profiles of responses to item options was r = .91, and the correlation for standard deviations of responses to item options was r = .66 (with 8 options for 30 items, or 240 observations, in each set). These correlations indicate that the responses across countries were about as similar as one could hope for, given the reliabilities of the data.

6.3.3 Item characteristics

We analyzed item characteristics for the combined samples using 1/0 scoring. The range of facilities (p-values) was from .42 to .84, with a mean p-value of .62. The range of difficulties (Δ values) was from 9.1 to 13.4 with a mean of 11.8. The range of discriminating power for items was computed with both biserial and point-biserial correlations. The range of r_{bis} was from .38 to .72, with a mean of .53. The range of r_{pbis} was from .29 to .53, with a mean of .41.

Specifics for the 1-Parameter Logistic (Rasch) and 2-Parameter Logistic (Birnbaum) Models. One- and two-parameter-logistic (PL) models, as derived from item-response theory (IRT), were fit to the data. The difference in the maximal marginal likelihoods was 84.7, with 30 degrees of freedom, indicating the superior fit of the two-parameter model. For the one-parameter model, 7 items had statistically significant chi-squares ($p < .05$), whereas for the two-parameter model, none of the items did.

6.3.4 Item-information statistics

Table 1 shows the point of maximum information for each of the 30 items in the ESJI. The results suggest that the items tended to be rather easy, on the whole, and that subsequent versions of the ESJI probably need some more difficult items.

Table 1

Item information statistics

Point of Maximum Information					
1.	- .5335	11.	- .3486	21.	-1.3415
2.	-1.8624	12.	- .2408	22.	.0778
3.	- .0595	13.	-1.1276	23.	- .8488
4.	- .1422	14.	-1.4116	24.	-1.1307
5.	- .3898	15.	-1.4472	25.	-1.3093
6.	-1.5271	16.	- .8430	26.	- .1914
7.	.0279	17.	- .6911	27.	- .6369
8.	-1.1277	18.	.1810	28.	- .6239
9.	-1.7692	19.	.2232	29.	- .7331
10.	- .2972	20.	.5614	30.	.3627

6.3.5 Factor analyses

Exploratory factor analysis. Exploratory principal-factor analysis was done, yielding 5 factors with eigenvalues greater than 1 in the Spanish sample (and .9 in the U.S. sample). These factors accounted for roughly 50% of the variance in the Spanish data and 63% of the variance in the U.S. data. Eigenvalues were 9.9, 1.5, 1.2, 1.2, and 1.1 respectively for the Spanish data and 14.3, 1.5, 1.2, 1.0, and .9 in respectively for the U.S. data.

Confirmatory factor analysis. Confirmatory factor analysis also was done. The conceptual framework proposed in Figure 1 provided a 3 x 5 model for the items of the ESJI. We therefore evaluated corresponding three- and five-factor latent variable models for the U.S. and Spanish data.

Initially, we evaluated 4 different models: (a) a 3-factor model for the U.S. sample, $\bullet_{(402)}^{2}$ = 918.4, RMSEA (root-mean-square error of approximation) = .08, CFI (comparative fit index) = .87, IFI (incremental fit index) =.87, and GFI (goodness of fit index) = .78; (b) a 3-factor model for the Spanish sample, $\bullet_{(402)}^{2}$ = 582.2, RMSEA = .043, CFI = .91, IFI = .91, and GFI = .86; (c) a 5-factor model for the USA sample, $\bullet_{(395)}^{2}$ = 878.6, RMSEA = .08, CFI = .88, IFI =.88, and GFI =.79; and (d) a 5-factor model for the Spanish sample, $\bullet_{(395)}^{2}$ = 526.0, RMSEA = .04, CFI = .94, IFI = .94, and GFI = .87. Two conclusions were drawn from these results. First, the model-fit indexes were comparable for the U.S. and Spanish samples, suggesting that the data could be combined in a single analysis. Second, overall, the fit indexes were better for the 5-factor model than for the 3-factor model, suggesting that the 5-factor structure was the preferred latent structure of the inventory. Because results were comparable for the U.S. and Spanish samples, we combined them in a single 5-factor multi-group model. We fitted four different modifications of this model: (a) a model equating the correlations between the latent variables in both samples, $\bullet_{(800)}^{2}$ = 1421.2, RMSEA = .06, CFI = .90, IFI = .90, and GFI = .87; (b) a model equating the correlations between the latent variables and measurement errors in both samples, $\bullet_{(830)}^{2}$ = 1643.6, RMSEA = .07, CFI = .87, IFI = .87, and GFI = .83; (c) a model equating the correlations between latent variables and factor loadings of the measured variables on the latent variables, $\bullet_{(830)}^{2}$ = 1627.6, RMSEA = .07, CFI = .87, IFI = .87,

and GFI = .83; and (d) a model equating the correlations between the latent variables, measurement errors, and factors loading of the observed variables on the latent variables, $\bullet_{(860)}^{2}$ = 1857.5, RMSEA = .08, CFI = .84, IFI = .84, and GFI = .80. According to these indexes, Model (a) described the data the best, suggesting that the underlying latent structure of the inventory is invariant across the U.S. and Spanish samples, but the measurement errors and factor loadings differ in the two samples.

The variables in the five-factor model were related. Table 2 shows the intercorrelations of the latent variables. As can be seen in the table, these latent-variable correlations are extremely high, suggesting that the five factors of the model are highly correlated and may represent a general factor, although not necessarily psychometric *g*, given that in past research subscales also have been highly correlated with each other but not with psychometric *g*.

Table 2

Latent variable correlations

	M	S	B	D	O
Motivating	1.00				
Situations	0.91	1.00			
	(0.02)				
	41.29				
Behaving	0.87	0.95	1.00		
	(0.02)	(0.02)			
	36.03	52.79			
Directions	0.84	0.94	0.96	1.00	
	(0.03)	(0.02)	(0.02)		
	32.25	50.09	56.77		
Organizing	0.87	0.95	0.98	0.96	1.00
	(0.02)	(0.02)	(0.02)	(0.02)	
	35.61	53.67	63.20	57.26	

6.4 External validation

6.4.1 Concurrent validities

To stress the consistency of the results across methods of scoring, we show r-based validity coefficients for the sample-specific data and 2-PL rescaled r-based validity coefficients for the combined sample data.

Supervisors' ratings. Concurrent validities for the rank-order correlation scoring of the ESJI are shown in Table 3 for the U.S. sample, the Spanish sample, and both samples combined. For the U.S. sample, validity coefficients ranged from .22 to .46 with a median of .36. All (8 of 12) correlations were statistically significant at the .05 level. For the Spanish sample, validity coefficients ranged from -.10 to .21 with a median of .14. Only three of the correlations were statistically significant. For the total sample, validity coefficients based on the 2-PL model ranged from .09 to .32 with a median of .20. Ten of the 12 correlations were statistically significant. Correlations were approximately the same for the d^2 scores and actually slightly better with the 1-PL model.

Table 3

Concurrent validities

Spanish sample (rhos)		
1.	My relationship with this employee is good:	-.10
2.	I think highly of this employee:	.10
3.	I am satisfied with this employee:	.05
4.	The employees relationships with other coworkers are good:	.06
1.	…Common sense ability:	.21 *
2.	…Academic ability:	.21 *
3.	…Creative ability:	.13
4.	…Working by him/herself:	.07
5.	…Working with others:	.14
6.	…Motivating him/herself:	.02
7.	…Managing tasks:	.16
8.	…Responsible:	.20 *
US sample (rhos)		
1.	My relationship with this employee is good:	.26 **
2.	I think highly of this employee:	.34 ***
3.	I am satisfied with this employee:	.38 ***
4.	The employees relationships with other coworkers are good:	.22 *
1.	…Common sense ability:	.42 ***
2.	…Academic ability:	.44 ***
3.	…Creative ability:	.37 ***
4.	…Working by him/herself:	.34 ***
5.	…Working with others:	.29 **
6.	…Motivating him/herself:	.46 ***
7.	…Managing tasks:	.41 ***
8.	…Responsible:	.30 **
Total sample (rescaled, 2-PL [Birnbaum] model)		
1.	My relationship with this employee is good:	.12
2.	I think highly of this employee:	.18 **
3.	I am satisfied with this employee:	.21 **
4.	The employees relationships with other coworkers are good:	.09
1.	…Common sense ability:	.32 ***
2.	…Academic ability:	.31 ***
3.	…Creative ability:	.20 **
4.	…Working by him/herself:	.17 *
5.	…Working with others:	.17 *
6.	…Motivating him/herself:	.18 **
7.	…Managing tasks:	.25 ***
8.	…Responsible:	.20 **

Note: Correlations are slightly better with 1PL

We cannot say for sure why the correlations for the U.S. sample were better than those for the Spanish sample. Some possible interpretations, based on what we know of the data, are (a) differential quality of supervisory ratings, (b) differential compositions of the two respective samples, and (c) relatively greater homogeneity of the Spanish sample relative to the U.S. sample. In particular, Spanish supervisors seemed less comfortable providing ratings of quality of performance than did U.S. supervisors, but because we did not anticipate this result, we collected no quantitative data with respect to it.

Self-ratings. The distributions of self-ratings proved not to be amenable to the validity analyses due to restriction of range in these variables. As has been found in other studies (e.g., Sternberg et al., 1981), participants tend to have almost a uniformly high opinion of themselves. With the item on planning career change deleted (because this item does not measure self-evaluation of competence), mean self-ratings on the 9-point scale ranged from 6.5 to 8.3, with an overall mean of 7.7 (sd = .47) for the United States sample and from 6.7 to 7.8 for the Spanish sample, with an overall mean of 7.1 (sd = .39). Corresponding standard deviations ranged from 1.0 to 2.4, with an overall mean standard deviation of 1.4 (sd = .40) for the United States sample. For the Spanish sample, standard deviations ranged from 1.1 to 1.9 with an overall mean of 1.5 (sd = .30). Ratings also were highly leptokurtic. Given the problem of restriction of range, we did not pursue further analyses of the self-ratings.

In sum, our main findings were of (a) satisfactory psychometric properties for the ESJI, (b) satisfactory to excellent internal-consistency reliabilities of the inventory, (c) excellent consistency of the data across cultures, (d) satisfactory internal validity using five-factor model, and (e) moderate concurrent validity in the U.S. sample and modest concurrent validity in the Spanish sample. We believe that the study showed the feasibility of our approach in measuring practical cognition.

7. Conclusions

Approximately 25 years ago, McClelland (1973) questioned the validity of cognitive-skill testing for predicting real-world criteria such as job performance, arguing in favor of competency tests that more closely would reflect job performance itself. Subsequent reviews of the literature on the predictive validity of cognition tests suggest that McClelland may have been pessimistic about the validity of cognition tests: individual differences in cognition-test performance account for, on average, between 4 and 25 percent of the variance in real-world criteria such as job performance (Barrett and Depinet, 1991; Hunter and Hunter, 1984; Schmidt and Hunter, 1998; Wigdor and Garner,1982). Nevertheless, these findings indicate that between 75 and 96 percent of the variance in real-world criteria such as job performance cannot be accounted for by individual differences in cognition-test scores. The emerging literature on practical cognition and similar constructs (e.g., social and emotional cognition) is a belated response to McClelland's call for new methods to assess practical skills. The literature and research reviewed in this volume provide several sources of evidence to support a distinction between academic and practical cognition.

First, the distinction between academic and practical cognition is entrenched in the conception of cognition held by laypeople and researchers alike. In addition to evidence provided by studies of implicit theories of cognition (e.g., Sternberg et al., 1981), analyses of researchers' descriptions of the nature of cognition suggest a prominent role for practical cognition. Seventy years ago, the editors of the *Journal of Educational Psychology* convened a symposium at which prominent psychological theorists of the day were asked to describe what they imagined cognition to be and what they considered the most crucial "next steps" in research. In a replication, Sternberg and Detterman (1986) posed these same questions to contemporary prominent theorists. An analysis of the responses of both cohorts of cognition theorists revealed concerns about practical aspects of cognition (Sternberg and Berg, 1986). For example, among the 42 crucial next steps that were mentioned by one or more theorists from either cohort, studying real-life manifestations of cognition was among the most frequently mentioned "next steps" of both the contemporary researchers and the original respondents. A distinction between academic and practical aspects of cognition also is supported by older adults' perception of age-related changes in their skill to think and solve problems (Williams, Denney, and Schadler, 1983). Three-fourths of the older adults sampled believed that their skill to solve practical problems increased over the years, despite the fact that performance on academic tasks begins to decline upon completion of formal schooling.

A second source of evidence to support a distinction between academic and practical cognition is the set of results of studies in which participants were assessed on both academic and practical tasks. These studies consistently find little or no correlation between performance on the two kinds of tasks. Overall, cognitive ability tests and similar measures are unrelated to (a) the order-filling performance of milk-processing plant workers (Scribner, 1986); (b) the degree to which racetrack handicappers employ a complex and effective algorithm (Ceci and Liker, 1986, 1988); (c) the complexity of strategies used in computer-simulated roles such as city manager (Dörner and Kreuzig, 1983; Dörner et al., 1983); and (d) the accuracy with which grocery shoppers identified quantities that provided the best value (Lave et al., 1984; Murtaugh, 1985). This research shows that the performance of both children and adults is susceptible to the context in which skills are measured. When problems are presented in a familiar context, whether that context is school or work, individuals appear more intelligent (e.g., Carraher et al., 1985; Roazzi, 1987).

A third source of support for the importance of practical skills comes from theories of managerial performance. Rational theories that are based on conventional notions of how people solve problems (e.g., Kepner and Tregoe, 1965; Plunkett and Hale, 1982) do not accurately represent the problem solving of experienced and successful managers. These observations led theorists to describe managerial problem solving as non-linear, convoluted, and action-oriented (e.g. McCall and Kaplan, 1985; Mintzberg et al., 1976). Furthermore, knowledge of how to solve problems can be characterized as tacit, and it may only enter into conscious awareness through reflection (Schön, 1983). The recognition that rational models of managerial problem solving do not explain the behavior of successful practitioners suggests that alternative approaches are needed to identify the practical skills underlying performance.

Finally, the research on tacit knowledge described throughout this volume offers an approach to understanding practical cognition. Over the course of studies with academic psychologist (Wagner, 1987; Wagner and Sternberg, 1985), business managers (Wagner and Sternberg, 1990), salespersons (Wagner et al., 1994), U.S. Air Force recruits (Eddy, 1988), and military leaders (Hedlund et al., 1999), we have found that tacit knowledge offers insight into the practical skills associated with success.

Several conclusions can be drawn from this program of research. First, these studies showed that tacit knowledge exists in the stories successful practitioners share about the lessons they learned in the process of performing their respective roles. These stories provide rich insights about the practically-oriented knowledge that practitioners are often unaware that they have acquired. Second, we showed that tacit knowledge can be measured through instruments that take into account the procedural and context-specific nature of tacit knowledge. Third, using such instruments, we have found that individuals who exhibit the skill to acquire and use tacit knowledge are more effective in their respective performance domains. Furthermore, tacit knowledge helps to explain some of the additional variance in performance that is not accounted for by measures of general cognitive skill. Fifth, although the acquisition of tacit knowledge may be influenced, to some extent, by *g* and amount of experience, tacit-knowledge inventories are not simply new measures of these constructs. Finally, tacit knowledge generally appears to be a singular construct within domains, but the content of tacit knowledge varies across domains. In other words, tacit knowledge appears to reflect a single underlying skill, which we label practical cognition. But, this underlying skill is not sufficient for performing well on domain-specific tacit-knowledge tests. Experience in a particular domain is important in the acquisition of tacit knowledge.

Based on consistent findings that tacit knowledge contributes to our understanding performance in a variety of domains, we discussed a number of potential ways to promote the acquisition and use of tacit knowledge. Numerous insights and products are obtained through the process of studying tacit knowledge. The categories of tacit knowledge within a domain, for example, offer insight into the experiences that provide important developmental opportunities. The products, such as the stories and the inventory questions, can be used to share the tacit knowledge with other practitioners. The tacit-knowledge research also suggests that training activities, such as case studies and simulations, may be valuable ways to impart experience-based, tacit knowledge and to provide opportunities to acquire new practical knowledge. Although these approaches may encourage the acquisition and use of tacit knowledge, in rapidly changing, complex environments, it may be more effective in the long run to identify and develop ways to help individuals to learn better from their everyday experiences.

Up to this point, our research efforts have been targeted primarily at understanding and measuring practical cognition. For the present and foreseeable future, we believe that the most viable approach to increasing the variance accounted for in real-world criteria such as job performance is to supplement existing cognition and aptitude tests with selection of additional measures based on new constructs such as practical cognition. Although we are excited by the promise of a new generation of measures of practical cognition, we are the first to admit that existing evidence for the new measures does not yet match that available for traditional cognitive-academic skill tests. However, a substantial amount of evidence indicates that performance on measures of practical cognition is related to a wide variety of criterion measures of real-world performance, but relatively unrelated to traditional measures of academic cognition. Consequently, using both kinds of measures explains more variance in performance than relying on either kind alone. Cognition is not only academic, but practical.

8. Future directions

There still are some issues that need to be addressed in further work on the scale, and we plan to address these issues in the next version.

1. Face validity. The use of the term *boss* proved to be a mistake, as it carries a somewhat negative connotation in European and other countries were hierarchical arrangements of workers are less socially acceptable than in the United States. The inventory also relied too much on office-type settings and needs to be expanded to include not only a broader range of occupational settings, but settings outside of the workplace. Our goal in the next version of our questionnaire, therefore, is to do extensive revision for face validity to ensure that the questionnaire we use will be viewed as face valid by all test-takers. Following a procedure we have used in other research, our plan is to ask test-takers to evaluate the face validity of the questionnaire by asking them how realistic the scenarios and solutions are.
2. Length. The ESJI was longer than would be ideal, given realistic constraints upon administration time. It could be and should be shortened both with respect to number of items (from 30 to, perhaps, 20) and number of items (from 8 to, perhaps, 5).
3. Description of test. The test originally was described as a test of practical ability or of everyday cognition. We have changed the name, effective as of this article, to reflect better both what the test measures and the constraints of the sociopolitical context in which the test will be administered. We thus will refer to the test as an everyday-situation-judgment test.
4. Number of scale points. The number of scale points were item will be seven in order to ensure that each point supplies useful information. Each scale point will have a verbal label describing its meaning.
5. Need for anchors in scoring. Scoring in the future will be converted to a 0-500 scale in order to match other assessments in the study of adult competencies in OECD countries. The mean will be 250 and the standard deviation, 50. We will anchor score points to specific levels of competency verbally described. Specific levels of competency will be anchored to various indicators of job performance, career satisfaction, and life satisfaction (with relevant data to be obtained from a new validity study).
6. Objective scoring. In addition to the type of prototype scoring we have used, we will experiment with right-wrong scoring based on sets of values that our previous research indicates seem to be universally accepted as indicating preferred behavior (e.g., honesty, sincerity, hard work, compassion, etc.)
7. Uninformative options. Options that elicit means very near the middle of the scale together with high standard deviations will be eliminated, as such options tend to be ones on which there is substantial disagreement among respondents.
8. Skills taxonomy. We already have begun development of a skills taxonomy and this development will be formalized for the next version of the ESJI. We have discovered that certain behaviors tend to be valued cross-situationally and other behaviors to be devalued. Examples of valued behaviors are productivity, honesty, politeness, serving as a good example to others, and doing what needs to be done even if one does not want to do it. Examples of devalued behaviors are lying, not working when one should, passing the buck, blaming others for one's own mistakes, failing to meet one's responsibilities, and doing things grudgingly.
9. Test-retest reliabilities. The present design did not allow for computation of test-retest reliabilities to measure scale stability, but future studies will do so.
10. Discriminant validities. We need a subsequent design that will permit us to assess the discriminant validity of the test with respect to *g*-based measures of abilities.
11. Range of skills tested. We need to develop a wider range of behaviors to be sampled, including behaviors from more varied jobs and behaviors that occur outside of jobs.

12. Translation. The procedures for translation were rather informal in this study and resulted in a few items that, in retrospect, were not ideally translated. The result was a lowering of the cross-sample correlation (which was nevertheless .91). A more careful procedure for ensuring the accuracy of translation is needed in a subsequent study, as well as for ensuring that the situations sampled adequately represent apposite situations in the two cultures assessed as well as others.
13. Focus groups. In subsequent work, we also plan prior to validation to have focus groups in each culture that will evaluate the relevance of each potential inventory situation for cultural appropriateness. Scenarios that are not deemed to be culturally valid by two-thirds of the focus group will be replaced.
14. Occupational groups. We will score by occupational groups in order to determine whether different occupational groups tend to prefer different responses to options.
15. Cultures. We would hope to introduce a third cultural context in subsequent work. Sternberg, Grigorenko, and Gil are all willing to participate in subsequent phases of this work.

In sum, the ESJI appears to be a promising inventory for the assessment of situational judgments, but further work is needed in order to refine the existing measure.

References

Ackerman, P. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

Ackerman, P. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D.K. Detterman (Ed.), *Current topics in human intelligence: Theories of cognition* (Vol. 4, pp. 1-27). Norwood, NJ: Ablex.

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

Aldwin, C.M., Sutton, K.J., Chiara, G., and Spiro, A. (1996). Age differences in stress, coping, and appraisal: Findings from the normative aging study. *Journal of Gerontology: Psychological Sciences, 51B,* 178-188.

Anastasi, A., and Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Anderson, J.R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J.R. (1986). Knowledge compilation: The general learning mechanism. In R.S. Michalski, J.G. Carbonell, and T.M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach.* Los Altos, CA: Kaufman.

Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192-210.

Anderson, J.R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Archer, D. (1980). *How to expand your social intelligence quotient*. New York: M. Evans.

Archer, D., and Akert, R.M. (1980). The encoding of meaning: A test of three theories of social interaction. *Sociological Inquiry, 50,* 393-419.

Argyris, C. (1991). Teaching smart people how to learn. *Harvard Business Review, 69,* 99-109.

Bäckman, L. and Dixon, R.A. (1992). Psychological compensation: A theoretical framework. *Psychological Bulletin, 112,* 259-283.

Baltes, P.B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611-626.

Baltes, P.B. (1993). The aging mind: Potentials and limits. *The Gerontologist, 33,* 580-594.

Baltes, P.B. (1997). On the incomplete architecture of human ontogeny. *American Psychologist, 52,* 366-380.

Baltes, P. (in press). *Wisdom: the orchestration of mind and virtue.* Boston: Blackwell.

Baltes, P.B. and Baltes, M.M. (1990). Psychological perspectives on successful aging: A model of selective optimization with compensation. In P.B. Baltes and M.M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences.* Cambridge: Cambridge University Press.

Baltes, P.B., Dittmann-Kohli, F., and Dixon, R.A. (1984). New perspectives on the development of intelligence in adulthood: Toward a dual-process conception and a model of selective optimization with compensation. In P.B. Baltes and O.G. Brim (Eds.), *Life-span development and behavior* (Vol 6., pp. 33-76). New York: Academic Press.

Baltes, P.B., and Lindenberger, U. (1988). On the range of cognitive plasticity in old age as a function of experience: 15 years of intervention research. *Behavior Therapy, 19,* 282-300.

Baltes, P.B., Sowarka, D., and Kliegl, R. (1989). Cognitive training research on fluid intelligence in old age: What can older adults achieve by themselves? *Psychology and Aging, 4,* 217-221.

Baltes, P.B., and Willis, S.L. (1982). Plasticity and enhancement of cognitive functioning in old age: Penn State's Adult Development and Enrichment Project (ADEPT). In F.I.M. Clark and S.E. Trehub (Eds.), *Aging and cognitive processes* (pp. 353-389). New York: Plenum.

Band, E. B., and Weisz, J.R. (1988). How to feel better when it feels bad: Children's perspective on coping with everyday stress. *Developmental Psychology, 24,* 247-253.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Barker, R.G. (Ed.) (1978). *Habitats, environments, and human behavior.* San Francisco: Jossey-Bass.

Barnes, M.L., and Sternberg, R.J. (1989) Social intelligence and decoding of nonverbal clues. Intelligence, *13,* 263-287.

Bar-On, R. (1997). *The Emotional Quotient Inventory (EQI): Technical Manual*. Toronto, Canada: Multi-Health Systems.

Barrett, G.V., and Depinet, R.L. (1991). A reconsideration of testing for competence rather than for intelligence. *American Psychologist, 46,* 1012-1024.

Barrett, G.V., Mihal, W.L., Panek, P.E., Sterns, H.L., and Alexander, R.A. (1977). Information processing abilities predictive of accident involvement for younger and older commercial drivers. *Industrial Gerontology, 4,* 173-182.

Barrett, G.V., and Watkins, S.K., (1986). Word familiarity and cardiovascular health as determinants of age-related recall differences. *Journal of Gerontology, 41,* 222-224.

Beck, A.T., Epstein, N., Brown, G., and Steer, R.A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Belmont, J.N., and Butterfield, E.C. (1969). The relations of short-term memory to development and cognition. In L. Lipsitt and H. Reese (Eds.), *Advances in child development and behavior* (Vol. 4., pp. 30-83). New York: Academic Press.

Bennett, G.K., Seashore, H.G., and Wesman, A.G. (1974). *The Differential Aptitude Tests* (Form T). New York: The Psychological Corporation.

Berg, C.A. (1984). Perspectives for viewing cognitive development throughout the life course. In R.J. Sternberg and C.A. Berg (Eds.), *Cognitive Development* (pp. 1-15). New York, NY: Cambridge University Press.

Berg, C.A. (1989). Knowledge of strategies for dealing with everyday problems from childhood through adolescence. *Developmental Psychology, 25,* 607-618.

Berg, C.A. (in press). The development of adult intelligence. In R.J. Sternberg (Ed.), *Handbook of intelligence*. New York, NY: Cambridge University Press.

Berg, C.A., and Calderone, K. (1994). The role of problem interpretations in understanding the development of everyday problem solving. In R.J. Sternberg and R.K. Wagner (Eds.), *Mind in context* (pp. 105-132). New York, NY: Cambridge University Press.

Berg, C.A., Calderone, K. and Gunderson, M. (1990, November). *Strategies young and old adults use to solve their own everyday problems.* Paper presented at the meeting of the Gerontological Society, Boston, MA.

Berg, C.A., Calderone, K., Sansone, C., Strough, J., and Weir, C. (1998). The role of problem definitions in understanding age and context effects on strategies for solving everyday problems. *Psychology and Aging, 13,* 29-44.

Berg, C.A., Hertzog, C., and Hunt, E. (1982). Age differences in the speed of mental rotation. *Developmental Psychology, 18,* 95-107.

Berg, C.A., Klaczynski, P. (1996). Practical intelligence and problem solving: Searching for perspective. In F. Blanchard-Fields and T.M. Hess (Eds.), *Perspectives on intelligence in adulthood and aging* (pp. 323-357). New York: McGraw-Hill.

Berg, C.A., Klaczynski, P., Calderone, K.S., and Strough, J. (1994). Adult age differences in cognitive strategies: Adaptive or deficient. In J. Sinnott (Ed.), *Interdisciplinary Handbook of Adult Lifespan Learning* (pp. 371-388). Westport, CT: Greenwood Press.

Berg, C.A., and Sternberg, R.J. (1985). A triarchic theory of cognitive development during adulthood. *Developmental Review, 5,* 334-370.

Bernadin, H.J., and Smith, P.C. (1981). A clarification of some issues regarding the development and use of behaviorally anchored rating scales (BARS). *Journal of Applied Psychology, 65,* 458-463.

Berry, J.W. (1974). Radical cultural relativism and the concept of intelligence. In J.W. Berry and P.R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 225-229). London: Methuen.

Berry, J.W. (1984). Towards a universal psychology of cognitive competence. In P.S. Fry (Ed.), *Changing conceptions of cognition and cognitive functioning* (pp. 35-61). Amsterdam: North-Holland.

Berry, J.W., and Irvine, S.H. (1986). Bricolage: Savages do it daily. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 271-306). New York: Cambridge University Press.

Bickhard, M.H. (1978). The nature of developmental stage. *Human Development, 21,* 217-233.

Blanchard-Fields, F. (1986). Reasoning and social dilemmas varying in emotional saliency: An adult developmental perspective. *Psychology and Aging, 1,* 325-333.

Blanchard-Fields, F. (1994). Age differences in causal attributions from an adult developmental perspective. *Journal of Gerontology: Psychological Sciences, 49,* 43-51.

Blanchard-Fields, F., Jahnke, H.C., and Camp, C. (1995). Age differences in problem-solving style: The role of emotional salience. *Psychology and Aging, 10,* 173-180.

Blanchard-Fields, F., and Norris, L. (1994). Causal attributions from adolescence through adulthood: Age differences, ego level, and generalized response style. *Aging Neuropsychology and Cognition, 1,* 67-86.

Bloom, B.S. (Ed.) (1985). *Developing talent in young people.* New York: Ballantine.

Bogen, J.E. (1975). Some educational aspects of hemispheric specialization. *UCLA Educator, 17,* 24-32.

Borman, W.C. (1991). Job behavior, performance, and effectiveness. In M.D. Dunnette and L.M. Hough (Eds.), *Handbook of industrial and organizational psychology, Vol. 2* (pp. 271-326). Palo Alto, CA: Consulting Psychologists Press, Inc.

Boyatzis, R.E. (1982). *The competent manager*. New York: John Wiley.

Brandtstaedter, J., and Greve, W. (1994). The aging self: Stabilizing and protective processes. *Developmental Review, 14,* 52-80.

Bray, D.W. (1982). The Assessment Center and the study of lives. *American Psychologist, 37,* 180-189.

Broadbent, D.E., and Aston, B. (1978). Human control of a simulated economic system. *Ergonomics, 21,* 1035-1043.

Broadbent, D.E., Fitzgerald, P., and Broadbent, M.H.P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology, 77,* 33-50.

Broca, P.P. (1861). Nouvelle observation d'aphemie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolutions frontales gauches. *Bulletins de la Societé Anotomique de Paris, 36,* 398-407.

Bronfenbrenner, U., and Ceci, S.J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review, 101,* 568-586.

Brown, A.L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 103-152). New York: Academic Press.

Brown, L.T., and Anthony, R.G. (1990). Continuing the search for social intelligence. *Personality and Individual Differences, 11,* 463-470.

Bruner, J.S., Shapiro, D., and Tagiuri, R. (1958). The meaning of traits in isolation and in combination. In R. Tagiuri and I. Petrollo (Eds.), *Person perception and interpersonal behavior* (pp. 277-288). Stanford, CA: Stanford University Press.

Burke, M.J., and Day, R.R. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology, 71,* 232-246.

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

Campbell, J.P. (1988). Training design for performance improvement. In J.P. Campbell, R.J. Campbell, and Associates (Eds.), *Productivity in organizations* (pp. 177-216). San Francisco: Jossey-Bass.

Cantor, N. (1978). Prototypicality and personality judgments. Unpublished doctoral dissertation, Stanford University.

Cantor, N. (1990). From thought to behavior: "Having" and "doing" in the study of personality and intelligence. *American Psychologist, 45,* 735-750.

Cantor, N., and Harlow, R. (1994). Social intelligence and personality: Flexible life-task pursuit. In R.J. Sternberg and P. Ruzgis (Eds.), *Personality and intelligence* (pp. 137-168). Cambridge, U.K.: Cambridge University Press.

Cantor, N. and Kihlstrom, J.F. (1987). *Personality and social intelligence.* Englewood Cliffs, N.J.: Prentice Hall.

Cantor, N., and Kihlstrom, J.F. (1989). Social intelligence and intellectual assessments of personality. In R.S. Wyer and T.K. Srull (Eds.), *Advances in Social Cognition* (Vol. 2, pp. 1-59). Hillsdale, N.J.: Erlbaum.

Cantor, N., Norem, J.K., Niedenthal, P.M., Langston, C.A., and Brower, A.M. (1987). Life tasks, self-concept ideals, and cognitive strategies in a life transition. *Journal of Personality and Social Psychology, 53,* 1178-1191.

Carraher, T.N., Carraher, D., and Schliemann, A.D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3,* 21-29.

Caruso, D.R., and Mayer, J.D. (1997). A quick scale of empathy. Unpublished manuscript.

Cattell, R.B. (1940). A culture free intelligence test. I. *Journal of Educational Psychology, 31,* 161-180.

Cattell, R.B. (1971). *Abilities: Their structure, growth and action.* Boston: Houghton Mifflin.

Cattell, R.B., and Cattell, H.E.P. (1963). *Test of g: Culture Fair, Scale 3.* Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R.B., and Cattell, H.E.P. (1973). *Measuring intelligence with the Culture Fair Tests.* Champaign, IL: Institute for Personality and Ability Testing.

Ceci, S.J. (1990). *On intelligence. . .more or less: A bio-ecological treatise on cognitive development.* Englewood Cliffs, NJ: Prentice-Hall.

Ceci, S.J., and Brofenbrenner, U. (1985). Don't forget to take the cupcakes out of the oven: Strategic time-monitoring, prospective memory and context. *Child Development, 56,* 175-190.

Ceci, S.J., and Liker, J. (1986). Academic and nonacademic intelligence: an experimental separation. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 119-142). New York: Cambridge University Press.

Ceci, S.J., and Liker, J. (1988). Stalking the IQ-expertise relationship: When the critics go fishing. *Journal of Experimental Psychology: General, 117,* 96-100.

Ceci, S.J. and Roazzi, A. (1994). The effects of context on intelligence: postcards from Brazil. In R.J. Sternberg and R.K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 74-101). New York: Cambridge University Press.

Ceci, S.J., and Ruiz, A. (1991). Cognitive complexity and generality: A case study. In R. Hoffman (Ed.), *The psychology of expertise.* New York: Springer-Verlag.

Chan, D., and Schmitt, N. (1998). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82,* 143-159.

Chapin, F.S. (1942). Preliminary standardization of a social impact scale. *American Sociological Review, 7,* 214-225.

Chapin, F.S. (1967). *The Social Insight Test.* Palo Alto, CA: Consulting Psychologists Press.

Charness, N. (1979). Components of skill in bridge. *Canadian Journal of Psychology, 33,* 1-16.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 467-476.

Charness, N. (1991). Expertise in chess: the balance between knowledge and search. In K.A. Ericsson and J. Smith (Eds.), *Toward a general theory of expertise* (pp. 39-63). New York: Cambridge University Press.

Charness, N., and Bieman-Coplan, S. (1994). The learning prospective: Adulthood. In R.J. Sternberg and C.A. Berg (Eds.), *Cognitive Development* (pp. 301-327). New York, NY: Cambridge University Press.

Charness, N., and Bosman, E.A. (1990). Expertise and aging: Life in the lab. In T.M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 343-385). Amsterdam: Elsevier Science.

Charness, N., and Bosman, E.A. (1995). Compensation through environmental modification. In R.A. Dixon and L. Baeckman, (Eds). *Compensating for psychological deficits and declines: Managing losses and promoting gains.* (pp. 147-168). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Charness, N., Krampe, R., and Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K.A. Ericsson (Ed.), *The road to excellence* (pp. 51-80). Hillsdale, NJ: Erlbaum.

Chase, W.G., and Ericsson, K.A. (1982). Skill and working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1-58). New York: Academic Press.

Chase, W.G., and Simon, H.A. (1973). The mind's eye in chess. In W.G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

Chen, M.J. (1994). Chinese and Australian concepts of intelligence. *Psychological Developmental Sociology, 6,* 101-117.

Chen, S.A., and Michael, W.B. (1993). First-order and higher-order factors of creative social intelligence within Guilford's structure-of-intellect model: a reanalysis of a Guilford data base. *Educational and Psychological Measurement, 53,* 619-641.

Chi, M.T.H., and Ceci, S.J. (1987). Content knowledge: Its role, representation, and restructuring in memory development. In H.W. Reese (Ed.), *Advances in child development and behavior, Vol. 20* (pp. 91-142). Orlando, FL: Academic Press.

Chi, M.T.H., Glaser, R., and Farr, M.J. (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Cohen, M. (1996). Preschoolers' practical thinking and problem solving: the acquisition of an optimal solution strategy. *Cognitive Development, 11,* 357-373.

Cohen, J., and Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colonia-Willner, R. (1998). Practical intelligence at work: Relationship between aging and cognitive efficiency among managers in a bank environment. *Psychology and Aging, 13,* 45-57.

Cornelius, S.W. (1984). Classic pattern of cognitive aging: Test familiarity, difficulty, and performance. *Journal of Gerontology, 39,* 201-206.

Cornelius, S.W., and Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging, 2,* 144-153.

Cosier, R.A., and Aplin, J.C. (1982). Intuition and decision making: some empirical evidence. *Psychological Reports, 51,* 275-281.

Cronbach, L.J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671-684.

Csikszentmihalyi, M. (1996). *Creativity*. New York: Harper Collins.

Das, J.P., and Naglieri, J.A. (1997). *Das-Naglieri cognitive assessment system.* Chicago, Illinois: Riverside Publishing Company.

Dasen, P. (1984). The cross-cultural study of cognition: Piaget and the Baoule. *International Journal of Psychology, 19,* 407-434.

Davidson, J.E., and Sternberg, R.J. (1984). The role of insight in cognitive giftedness. *Gifted Child Quarterly, 28,* 58-64.

Davidson, J.E., and Sternberg, R.J. (1998). Smart problem solving: How metacognition helps. In D.J. Hacker, A.C. Graesser, and J. Dunlosky (Eds.), *Metacognition in educational theory and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Davies, J., and Easterby-Smith, M. (1984). Learning and developing from managerial work experiences. *Journal of Management Studies, 2,* 169-183.

Davies, M., Stankov, L., and Roberts, R.D. (1998). What remains of emotional intelligence? *Journal of Personality and Social Psychology: Attitudes and Social Cognition, 75,* 989-1015.

De Groot, A. (1978). *Thought and choice in chess.* The Hague: Mouton. (Original work published 1946.)

Denney, N.W. (1979). Problem solving in late life: Intervention research. In P.B. Baltes and O.G. Brim (Eds.), *Life-span development and behavior* (Vol. 2., pp. 37-66). New York: Academic Press.

Denney, N.W. (1982). Aging and cognitive changes. In B.B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 807-827). Englewood Cliffs, NJ: Prentice-Hall.

Denney, N.W. (1989). Everyday problem solving: Methodological issues, research findings, and a model. In I.W. Poon, D.C. Rubin, and B.A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 330-351). New York, NY: Cambridge University Press.

Denney, N.W., and Palmer, A.M. (1981). Adult age differences on traditional and practical problem-solving measures. *Journal of Gerontology, 36,* 323-328.

Denney, N.W., and Pearce, K.A. (1989). A developmental study of practical problem sovling in adults. *Psychology and Aging, 4,* 438-442.

Denney, N.W., Pearce, K.A., and Palmer, A.M. (1982). A developmental study of adults' performance on traditional and practical problem-solving tasks. *Experimental Aging Research, 8,* 115-118.

Dixon, R.A. (1994). Contextual approaches to adult cognitive development. In R.J. Sternberg and C.A. Berg (Eds.), *Cognitive Development* (pp. 350-380). New York, NY: Cambridge University Press.

Dixon, R.A., and Baltes, P.B. (1986). Toward life-span research on the functions and pragmatics of intelligence. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 203-235). New York: Cambridge University Press.

Dodge, K.A., Pettit, G.S., McClaskey, C.L., and Brown, M.M. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 51,* 1-85.

Dörner, D., and Kreuzig, H. (1983). Problemlosefahigkeit und intelligenz. *Psychologische Rundschaus, 34,* 185-192.

Dörner, D., Kreuzig, H., Reither, F., and Staudel, T. (1983). *Lohhausen: Vom Umgang mit Unbestimmtheir und Komplexitat.* Bern: Huber.

Dreary, I., and Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist, 51,* 599-608.

Eddy, A.S. (1988). *The relationship between the Tacit Knowledge Inventory for Managers and the Armed Services Vocational Aptitude Battery*. Unpublished master's thesis, St. Mary's University, San Antonio, TX.

Egan, D.E., and Schwartz, B.J. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition, 7,* 149-158.

Engle, R.W., and Bukstel, L. (1978). Memory processes among bridge players of differing expertise. *American Journal of Psychology, 91,* 673-679.

Ericsson, K.A. (1996). The acquisition of expert performance. In K.A. Ericsson (Ed.), *The road to excellence* (pp. 1-50). Hillsdale, NJ: Erlbaum.

Ericsson, K.A., Krampe, R.T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363-406.

Ericsson, K.A., and Smith, J. (1991). Empirical study of expertise: Prospects and limits. In K.A. Ericsson and J.A. Smith (Eds.), *Toward a general theory of expertise* (pp. 1-38). New York: Cambridge University Press.

Farah, M.J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95,* 307-317.

Finke, R.A., Ward, T.B., and Smith, S.M. (1992). *Creative cognition: theory, research, and applications.* Boston, MA: MIT Press.

Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin, 51,* 327-358.

Flavell, J.H. (1970). Developmental studies of mediated memory. In H.W.Reese and L.P. Lipsitt (Eds.), *Advances in child development and child behavior* (Vol 5., pp. 181-211). New York: Academic Press.

Fleeson, W., and Cantor, N. (1995). Goal relevance and the affective experience of daily life: Ruling out situation explanation. *Motivation and Emotion, 19,* 25-57.

Folkman, S., Lazarus, R.S., Pimley, S., and Novacek, J. (1987). Age differences in stress and coping processes. *Psychology and Aging, 2,* 171-184.

Ford, D.H. (1987). *Humans as self-constructing living systems: A developmental perspective on behavior and personality*. Hillsdale, NJ: Erlbaum.

Ford, D.H. (1994). *Humans as self-constructing living systems: A developmental perspective on behavior and personality* (2nd ed.). State College, PA: Ideals.

Ford, M.E. (1982). Social intelligence and social competence in adolescence. *Developmental Psychology, 18,* 323-340.

Ford, M.E. (1986). For all practical purposes: Criteria for defining and evaluating practical intelligence. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 183-200). New York, NY: Cambridge University Press.

Ford, M.E., and Ford, D.H. (1987). *Humans as self-constructing living systems: Putting the framework to work.* Hillsdale, NJ: Erlbaum.

Ford, M.E., and Maher, M.A. (1998). Self-awareness and social intelligence. In M.D. Ferrari and R.J. Sternberg (Eds.), *Self-awareness: Its nature and development* (pp 191-218). New York: The Guilford Press.

Ford, M.E., and Tisak, M.S. (1983). A further search for social intelligence. *Journal of Educational Psychology, 75,* 196-206.

Frederiksen, N. (1966). Validation of a simulation technique. *Organizational Behavior and Human Performance, 1,* 87-109.

Frederiksen, N. (1986). Toward a broader conception of human intelligence. *American Psychology, 41,* 445-452.

Frederiksen, N., Carlson, S., and Ward, W.C. (1984). The place of social intelligence in a taxonomy of cognitive abilities. *Intelligence, 8,* 315-337.

Frederiksen, N., Jensen, O., and Beaton, A.E. (1972). *Prediction of organizational behavior.* New York: Pergamon Press.

Frederiksen, N., Saunders, D.R., and Wand, B. (1957). The in-basket test. *Psychological Monographs, 71*.

Freeman, N.H., Lewis, C., and Doherty, M.J. (1991). Preschoolers' grasp of a desire for knowledge in false-belief prediction: Practical cognition and verbal report. *British Journal of Developmental Psychology, 9,* 139-157.

French, J., and Raven, B.H. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies of social power*. Ann Arbor, MI: Institute for Social Research.

Galton, F. (1883). *Inquiry into human faculty and its development.* London: Macmillan.

Gardiner, J.M. (1988). Functional aspects of recollective experience. *Memory and Cognition, 16,* 309-313.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic.

Gardner, H. (1993). Seven creators of the modern era. In J. Brockman (Ed.), *Creativity* (pp. 28-47). New York: Simon and Schuster.

Gardner, H. (1998). Are there additional intelligences? The case for naturalist, spiritual, and existential intelligences. In J. Kane (Ed.), *Education, information, and transformation*. Englewood Cliffs, NJ: Prentice-Hall.

Gardner, H. (1999). Who owns intelligence? *The Atlantic Monthly, 283,* 67-76.

Gardner, H., Krechevsky, M., Sternberg, R.J., and Okagaki, L. (1994). Intelligence in context: Enhancing students' practical intelligence for school. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 105-127). Cambridge, MA: Bradford Books.

Gazzaniga, M.S. (1985). *The social brain: Discovering the networks of the mind.* New York: Basic Books.

Getzels, J., and Csikszentmihalyi, M. (1976). *The creative vision: A longitudinal study of problemfinding in art.* New York: WileyInterscience.

Gill, R., and Keats, D.M. (1980). Elements of cognitive competence: Judgments by Australian and Malay university students. *Journal of Cross-Cultural Psychology, 11,* 233-243.

Goleman, D. (1995). *Emotional intelligence.* New York: Bantam Books.

Goodnow, J.J. (1986). Some lifelong everyday forms of intelligent behavior: Organizing and reorganizing. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence* (pp. 31-50). New York: Cambridge University Press.

Gough, H.G. (1966). Appraisal of social maturity by means of the CPI. *Journal of Abnormal Psychology, 71,* 189-195.

Greenspan, S. (1981). Defining childhood social competence: A proposed working model. In B.K. Keogh (Ed.), *Advances in special education* (Vol. 3, pp. 1-39). Greenwich, CT: JAI Press.

Greenspan, S., and Driscoll, J. (1997). The role of intelligence in a broad model of personal competence. In D.P. Flanagan and J.L. Genshaft (Eds.), *Contemporary cognitive assessment: Theories, tests, and issues* (pp. 131-150). New York: Guilford Press.

Greenspan, S., and Granfield, J.M. (1992). Reconsidering the construct of mental retardation: Implications of a model of social competence. *American Journal on Mental Retardation, 96,* 442-453.

Grigorenko, E.L., Geissler, P.W., Prince, R., Okatcha, F., Nokes, C., Kenny, D.A., Bundy, D.A., and Sternberg, R.J. (in press). The organization of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development*.

Grigorenko, E.L., and Sternberg, R.J. (1998). Dynamic testing. *Psychological Bulletin, 124,* 75-11.

Grigorenko, E.L., and Sternberg, R.J. (in press). Analytical, creative, and practical intelligence as predictors of self-reported adaptive functioning: A case study in Russia. *Intelligence.*

Groen, G.J., and Patel, V.L. (1988). The relationship between comprehension and reasoning in medical expertise. In M.T. H. Chi, R. Glaser, and M. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Guilford, J.P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Guilford, J.P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review, 89,* 48-59.

Gustafsson, J.E. (1988). Hierarchical models of individual differences in cognitive abilities. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 35-71). Hillsdale, NJ: Erlbaum.

Haier, R.J., Nuechterlein, K.H., Hazlett, E., Wu, J.C., Pack, J., Browning, H.L., and Buchsbaum, M.S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence, 12,* 199-217.

Haier, R.J., Siegel, B., Tang, C., Abel, L., and Buchsbaum, M.S. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence, 16,* 415-426.

Halberstadt, A.G., and Hall, J.A. (1980). Who's getting the message?: Children's nonverbal abilities and their evaluation by teachers. *Developmental Psychology, 16,* 564-573.

Harter, S. (1983). Developmental prospectives on the self-system. In P.H. Mussen (Ed.), *Handbook of child psychology, Vol. IV.* New York, NY: Wiley.

Hartigan, J.A. (1975). *Clustering algorithms.* New York: Wiley.

Hartley, A.A. (1989). The cognitive etiology of problem solving. In L.W. Poon, D.C. Rubin, and B.A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 300-329). New York, NY: Cambridge University Press.

Havinghurst, R. (1972). *Developmental tasks and education*. New York: Van Nostrand.

Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. (1983). An overview of an expert system. In F. Hayes-Roth, D.A. Watterman, and D.B. Lenat (Eds.), *Building expert systems.* Reading, MA: Addison-Wesley.

Hayslip, B., Jr. (1989a). Alternative mechanisms for improvements in fluid ability performance among older adults. *Psychology and Aging, 4,* 122-124.

Hayslip, B., Jr. (1989b). Fluid ability training with aged people: A past with a future? *Educational Gerontology, 15,* 573-595.

Hebb, D.O. (1949). *The organization of behavior: A neuropsychological theory.* New York: Wiley.

Heckhausen, J., and Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102,* 284-304.

Hedlund, J., Forsythe, G.B., Horvath, J.A., Williams, W.M., Snook, S., Dennis, M., and Sternberg, R.J. (1999). Practical intelligence: The role of tacit knowledge in understanding leadership. Manuscript submitted for publication.

Hedlund, J., Horvath, J.A., Forsythe, G.B., Snook, S., Williams, W.M., Bullis, R.C., Dennis, M., and Sternberg, R.J. (1998). *Tacit Knowledge in Military Leadership: Evidence of Construct Validity* (Tech. Rep. 1080). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Hendricks, M., Guilford, J.P., and Hoepfner, R. (1969) Measuring creative social intelligence. *Reports from the Psychological Laboratory, University of Southern California,* No.42.

Hendrickson, A.E., and Hendrickson, D.E. (1980). The biological basis for individual differences in intelligence. *Personality and Individual Differences, I,* 3-33.

Hofland, B.F., Willis, S.L., and Baltes, P.B. (1981). Fluid intelligent performance in the elderly: Intraindividual variability and conditions of assessment. *Journal of Educational Psychology, 73,* 573-586.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33,* 307-316.

Horn, J.L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F.I.M. Craik and A. Trehum (Eds.), *Aging and cognitive processes* (pp. 237-278). New York: Plenum.

Horn, J.L. (1994). Theory of fluid and crystallized intelligence. In R.J. Sternberg (Ed.), *The encyclopedia of human intelligence* (Vol. 1, pp. 443-451). New York: Macmillan.

Horn, J.L., and Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57,* 253-270.

House, R.J. (1971). A path-goal theory of leadership effectiveness. *Administrative Science Quarterly, 16,* 321-339.

Howell, W.C., and Cooke, N.J. (1989). Training the human information processor: A review of cognitive models. In I.L. Goldstein (Ed.), *Training and development in organizations* (pp. 121-182). San Francisco: Jossey-Bass.

Hoyer, W.J., Labouview, G.V., and Baltes, P.B. (1973). Modification of response and speed deficits and cognitive performance in the elderly. *Human Development, 16,* 233-242.

Hunt, T. (1928). The measurement of social intelligence. *Journal of Applied Psychology, 12,* 317-334.

Hunt, E.B. (1978). Mechanics of verbal ability. *Psychological Review, 85,* 109-130.

Hunt, E., Frost, N., and Lunneborg, C. (1973). Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7, pp. 87-122). New York: Academic Press.

Hunt, E.B., Lunneborg, C., and Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology, 7,* 194-227.

Hunter, J.E., and Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Ilgen, D.R., Barnes-Farrell, J.L., and McKellin, D.B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes, 54,* 321-368.

Jacoby, L.L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 21-38.

Jensen, A.R. (1980*). Bias in mental testing.* New York: Free Press.

Jensen, A.R. (1982). Reaction time and psychometric g. In H.J. Eysenck (Ed.), *A model for intelligence.* Heidelberg: Springer-Verlag.

Jensen, A.R. (1993). Test validity: *g* versus "tacit knowledge," *Current Directions in Psychological Science, 1,* 9-10.

Jones, K. and Day, J.D. (1997). Discrimination of two aspects of cognitive-social intelligence from academic intelligence. *Journal of Educational Psychology, 89,* 486-497.

Kanfer, R., and Ackerman, P.L. (1989). Motivation and cognitive skills: An integrative aptitude—treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74,* 657-690.

Karmiloff-Smith, A. (1988). The child is a theorist not an inductivist. *Mind and Language, 3,* 183-196.

Kaufman, A.S., and Kaufman, N.L. (1993). *Kaufman Adolescent and Adult Intelligence Test.* Circle Pines, MN: American Guidance Service.

Keating, D.K. (1978). A search for social intelligence. *Journal of Educational Psychology, 70,* 218-233.

Kepner, C.H., and Tregoe, B.B. (1965). *The rational manager: A systematic approach to problem solving and decision making.* New York: McGraw-Hill.

Kerr, S., and Jermier, J.M. (1978). Substitutes for leadership: Their meaning and measurement. *Organizational Behavior and Human Performance, 22,* 375-403.

Keys, B., and Wolfe, J. (1990). The role of management games and simulations in education and research. *Journal of Management, 16,* 307-336.

Kihlstrom, J.F., and Cantor, N. (in press). Social intelligence. In R.J. Sternberg (Ed.), *Handbook of intelligence,* (2nd ed.). Cambridge, UK: Cambridge University Press.

Klaczynski, P.A., Laipple, J.S., and Jurden, F.H. (1992). Educational context differences in practical problem solving during adolescence. *Merrill-Palmer Quarterly, 38,* 417-438.

Kohlberg, L. (1963). The development of children's orientations toward a moral order: I. Sequence in the development of moral thought. *Vita Humana, 6,* 11-33.

Kohn, M.L., and Schooler, C. (Eds.) (1983). *Work and personality*. Norwood, NJ: Ablex.

Kosmitzki, C., and John, O.P. (1993). The implicit use of explicit conceptions of social intelligence. *Personality and Individual Differences, 15,* 11-23.

Kreitler, S., and Kreitler, H. (1987). Conceptions and processes of planning: The developmental perspective. In S.L. Friedman and E.K. Scholnick (Eds.), *Blueprints for thinking: The role of planning in cognitive development* (pp. 205-272). Cambridge, UK: Cambridge University Press.

Kuhn, D., Pennington, N., and Leadbeater, B. (1983). Adult thinking in developmental perspective. In P.B. Baltes and O.G. Brim (Eds.), *Life-span development and behavior* (Vol. 5). New York: Academic Press.

Laboratory of Comparative Human Cognition (1982). Culture and intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 642-719). New York: Cambridge University Press.

Labouvie-Vief, G. (1982). Dynamic development and mature autonomy. *Human Development, 25,* 161-191.

Labouvie-Vief, G. (1992). A Neo-Piagetian perspective on adult cognitive development. In R.J. Sternberg and C.A. Berg (Eds.), *Cognitive development* (pp. 197-228). New York, NY: Cambridge University Press.

Laipple, J.S. (1992). Problem-solving in young and old adulthood: The role of task interpretation. *Dissertation Abstracts International, 53(1-B),* 582.

Landy, F.J., and Farr, J. (1980). Performance ratings. *Psychological Bulletin, 87,* 72-107.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* New York: Cambridge University Press.

Lave, J., Murtaugh, M., and de la Roche, O. (1984). The dialectic of arithmetic in grocery shopping. In B. Rogoff and J. Lace (Eds.), *Everyday cognition: Its development in social context* (pp. 67-94). Cambridge, MA: Harvard University Press.

Lazarus, R.S., and Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

Legree, P.J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence, 21,* 247-266.

Lesgold, A., Rubinson, H. Feltovich, P., Glaser, R., Klopfer, D., and Wang, Y. (1985). *Expertise in a complex skill: Diagnosing X-ray pictures.* Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh Technical Report.

Likert, R. (1967). The human organization: Its management and value. New York: McGraw-Hill.

Loewen, E.R., Shaw, J.R., and Craik, F.I.M. (1990). Age differences in components of metamemory. *Experimental Aging Research, 16,* 43-48.

Manz, C.C., and Sims, H.P. Jr. (1980). Self-management as a substitute for leadership: A social learning theory perspective. *Academy of Management Review, 5,* 361-367.

Marlowe, H.A. (1986). Social intelligence: Evidence for multidimensionality and construct independence. *Journal of Educational Psychology, 78,* 52-58.

Mayer, J.D., Caruso, D., and Salovey, P. (in press). Emotional intelligence meets traditional standards for an intelligence. test. *Intelligence.*

Mayer, J.D., and Greeno, J.G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology, 63,* 165-173.

Mayer, J.D., and Salovey, P. (1993). The intelligence of emotional intelligence. Intelligence, *17,* 433-442.

Mayer, J.D., and Salovey, P. (1997). What is emotional intelligence? In P. Salovey and D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3-31). New York: Basic Books.

Mayer, J.D., Salovey, P., and Caruso, D. (in press). Competing models of emotional intelligence. In R.J. Sternberg (Ed)., *Handbook of intelligence* (2nd ed.). New York: Cambridge University Press.

McCall, M.W., and Kaplan, R.E. (1985). *Whatever it takes: Decision makers at work.* Englewood Cliffs, NJ: Prentice-Hall.

McClelland, D.C. (1973). Testing for competence rather than for "intelligence." *American Psychologist, 28,* 1-14.

McClelland, D.C. (1976). *A guide to job competency assessment.* Boston: McBer.

Meacham, J.A. and Emont, N.C. (1989). The interpersonal basis of everyday problem solving. In J.D. Sinnott (Ed.), *Everyday problem solving* (pp. 7-23). New York, NY: Praeger.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50 (9),* 741-750.

Miller, K.A., and Kohn, M.L. (1983). The reciprocal effects on job conditions and the cognitivity of leisure0time activities. In M.L. Kohn and C. Schooler (Eds.), *Work and personality* (pp. 217-241). Norwood, NJ: Ablex.

Minsky, M. (1968). *Semantic information processing.* Cambridge, MA: MIT Press.

Mintzberg, H., Raisinhani, D., and Theoret, A. (1976). The structure of "unstructured" decision processes. *Administration Science Quarterly, 21,* 246-275.

Mischel, W. (1984). Convergences and challenges in the search for consistency. *American Psychologist, 39,* 351-364.

Moss, F.A., and Hunt, T. (1927) Are you socially intelligent? *Scientific American, 137,* 108-110.

Moss, F.A., Hunt, T., Omwake, K.T., and Woodward, L.G. (1949). *Social Intelligence Test, George Washington University Series.* Washington, DC: Center for Psychological Services, George Washington University.

Motowidlo, S.J., Dunnette, M.D., and Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75,* 640-647.

Murtaugh, M. (1985). The practice of arithmetic by American grocery shoppers. *Anthropology and Education Quarterly, 16,* 186-192.

Naus, M.J. and Ornstein, P.A. (1983). Development of memory strategies: Analysis, questions and issues. In M.T.M. Chi (Ed.), *Trends in memory development research: Contributions to human development* (Vol. 9, pp. 1-30). Basel, Switzerland: S. Karger.

Neely, A.S., and Backman, L. (1993). Long-term maintenance of gains from memory training in older adults: Two 3 ½-year follow-up studies. *Journal of Gerontology, 48,* 233-237.

Neisser, U. (1976). General, academic, and artificial intelligence. In L. Resnick (Ed.), *Human intelligence: Perspectives on its theory and measurement* (pp. 179-189). Norwood, NJ: Ablex.

Neisser, U. (1979). The concept of intelligence. In R.J. Sternberg and D.K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 179-189). Norwood, NJ: Ablex.

Nelson, M.J., and Lamke, T.A. (1973). *The Henmon-Nelson test of mental ability*. Boston: Houghton Mifflin.

Nettlebeck, T. (1982). Inspection time: An index for intelligence. *Quarterly Journal of Experimental Psychology, 34,* 299-312.

Neugarten, B.L., Moore, J.W., and Lowe, J.C. (1968). Age norms, age constraints, and adult socialization. In B.L. Neugarten (Ed.), *Middle age and aging* (pp. 22-28). Chicago: University of Chicago Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., and Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Nuñes, T. (1994). Street intelligence. In R.J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp 1045-1049). New York: Macmillan.

Nuñes, T., Schliemann, A.D., and Carraher, D.W. (1993). *Street mathematics and school mathematics.* New York: Cambridge University Press.

Nunally, J.C. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill.

Okagaki, L., and Sternberg, R.J. (1993). Parental beliefs and children's school performance. *Child Development, 64(1),* 36-56.

Oltman, P.K., Raskin, F., and Witkin, H.A. (1971). *Group Embedded Figures Test.* Palo Alto, CA: Consulting Psychologists Press.

O'Sullivan, M., and Guilford, J.P. (1976). *Four factor tests of social intelligence (behavioral intelligence): Manual of instructions and interpretations.* Orange, CA: Sheridan Psychological Services.

O'Sullivan, M., Guilford, J.P., and deMille, R. (1965). The measurement of social intelligence. *Reports from the Psychological Laboratory, University of Southern California,* No. 34.

Patel, V.L., and Groen, G.J. (1991). The general and specific nature of medical expertise: a critical look. In K.A. Ericsson and J. Smith (Eds.), *Toward a general theory of expertise* (pp. 93-125). Hillsdale, NJ: Erlbaum.

Perret-Clermont, A.N. (1980). *Social interaction and cognitive development in children.* London: Academic Press.

Piaget, J. (1972). *The psychology of intelligence.* Totowa, NJ: Littlefield Adams.

Plunkett, L.C., and Hale, G.A. (1982). *The proactive manager.* New York: John Wiley and Sons.

Polanyi, M. (1966). *The tacit dimensions.* Garden City, NY: Doubleday.

Polanyi, M. (1976). Tacit knowledge. In M. Marx and F. Goodson (Eds.), *Theories in contemporary psychology* (pp. 330-344). New York: Macmillan.

Pressley, M., Forrest-Presley, D.L., Elliot-Faust, D., and Miller, G. (1985). Children's use of cognitive strategies: How to teach strategies, and what to do if they can't be taught. In M. Pressley and C.J. Brainers (Eds.), *Cognitive learning and memory in children: Progress in cognitive development research* (pp. 1-47). New York: Springer.

Raven, J.C. (1958). *Guide to using the Coloured Progressive Matrices*. London, England: H.K. Lewis and Co.

Raven, J.C., Court, J.H., and Raven, J. (1992). *Manual for Raven's Progressive Matrices and Mill Hill Vocabulary Scales.* Oxford: Oxford Psychologists Press.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 317-327.

Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81,* 115-119.

Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118,* 219-235.

Reber, A.S., and Millward, R.B. (1968). Event observation in probability learning. *Journal of Experimental Psychology, 77,* 317-327.

Ree, M.J. and Earles, J.A. (1993). *g* is to psychology what carbon is to chemistry: A reply to Sternberg and Wagner, McClelland, and Calfee, *Current Directions in Psychological Science, 1,* 11-12.

Reed, T.E. and Jensen, A.R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. *Intelligence, 16,* 259-272.

Reichert, U., and Dorner, D. (1988). Heurismen beim Umgang mit einem "einfachen" dynamischen System [Heuristics in the control of a "simple" dynamic system]. *Sprache and Kognition, 7,* 12-24.

Reitman, J. (1976). Skilled perception in GO: Deducing memory structures from interresponse times. *Cognitive Psychology, 8,* 336-356.

Rest, J. (1975). Longitudinal study of the Defining Issues Test of moral judgement: A strategy for analyzing developmental change. *Developmental Psychology, 11,* 738-748.

Riggio, R.E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51,* 649-660.

Riggio, R.E. (1989). *Manual for the Social Skills Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Riggio, R.E., Messamer, J., and Throckmorton, B. (1991). Social and academic intelligence: Conceptually distinct but overlapping constructs. *Personality and Individual Differences, 12,* 695-702.

Roazzi, A. (1987). Effects of context on cognitive development. In J.F. Cruz and R.A. Goncalves (Eds.), *Psicologia e Educao: Investigacao e intervencao*. Porto: Associacao dos Piscologos Portugueses.

Rogoff, B. (1982). Integrating context and cognitive development. In M.E. Lamb and A.L. Brown (Eds.), *Advances in development psychology* (Vol. 2, pp. 125-169). Hillsdale, NJ: Erlbaum.

Rogoff, B., Gauvain, M., and Gardner, W. (1987). Children's adjustment of plans to circumstances. In S.L. Friedman, E.K., Scholnick, and R.R. Cocking (Eds.), *Blueprints for thinking* (pp. 303-320). New York, NY: Cambridge University Press.

Rogoff, B. and Lave, J. (Eds.). (1984). *Everyday cognition: its development in social context.* Cambridge, Massachusetts: Harvard University Press.

Rosenthal, R. (Ed.) (1979). *Skill in nonverbal communication: Individual differences.* Cambridge, MA: Oelgeschlager, Gunn, and Hain.

Rosenthal, R., Hall, J.A., DiMatteo, M.R., Rogers, P.L., and Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS test.* Baltimore: Johns Hopkins University Press.

Ross, L.D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10). New York: Academic Press.

Salovey, P., and Mayer, J.D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185-211.

Salovey, P., and Mayer, J.D. (1994). Some final thoughts about personality and intelligence. In R.J. Sternberg, and P. Ruzgis (Eds.), *Personality and intelligence* (pp. 303-318). Cambridge, UK: Cambridge University Press.

Salthouse, T.A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General, 113,* 345-371.

Salthouse, T.A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, NJ: Erlbaum.

Salthouse, T.A. (1998). Relation of successive percentiles of reaction time distributions to cognitive variables and adult age. *Intelligence, 26,* 153-166.

Salthouse, T.A., and Somberg, B.L. (1982). Skilled performance: The effects of adult age and experience on elementary processes. *Journal of Experimental Psychology: General, 111,* 176-207.

Sansone, C., and Berg, C.A. (1993). Adapting to the environment across the life span: Different process or different inputs? *International Journal of Behavioral Development, 16,* 215-241.

Schacter, D.L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology, 13,* 501-518.

Schaie, K.W. (1977/1978). Toward a stage theory of adult cognitive development. *International Journal of Aging and Human Development, 8,* 129-138.

Schaie, K.W. (1986). Beyond calendar definitions of age, time, and cohort: The general developmental model revisited. *Developmental Review, 6,* 252-277.

Schaie, K.W. (1988). Variability in cognitive functioning in the elderly: Implications for societal participation. In A.D. Woodhead, M.A. Bender, and R.C. Leonard (Eds.), *Phenotypic variation in populations: Relevance to risk assessment* (pp. 191-212). New York: Plenum.

Schaie, K.W. (1989). The hazards of cognitive aging. *Gerontologist, 29,* 484-493.

Schaie, K.W. (1994). The course of adult development. *American Psychologist, 49,* 304-313.

Schaie, K.W. (1996). *Cognitive development in adulthood: The Seattle Longitudinal Study*. New York, NY: Cambridge University Press.

Schaie, K.W. and Willis, S.L. (1986). Can decline in adult cognitive functioning be reversed? *Developmental Psychology, 22,* 223-232.

Schank, R.C. (1990). *Tell me a story: A new look at real and artificial memory*. New York: Charles Scribner's Sons.

Schmidt, F.L., and Hunter, J.E. (1993). Tacit knowledge, practical intelligence, general mental ability, and job knowledge, *Current Directions in Psychological Science, 1,* 8-9.

Schmidt, F.L, and Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262-274.

Schooler, C. (1984). Psychological effects of complex environments during the life span: A review and theory. *Intelligence, 8,* 259-281.

Schooler, C. (in press). The cognitive effects of the demands of the work environment. In R.S. Sternberg and E.L. Grigorenko (Eds.) *Environmental effects on cognitive functioning*. Hillsdale, NJ: Lawrence Erlbaum.

Schon, D.A. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J., and Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25,* 167-177.

Scribner, S. (1984). Studying working cognition. In B. Rogoff and J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 9-40). Cambridge, MA: Harvard University Press.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 13-30). New York: Cambridge University Press.

Scribner, S., and Cole, M. (1981). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Serpell, R. (1982). Measures of perception, abilities, and intelligence. In W.W. Hartup (Ed.), *Review of Child Development Research.* Chicago: University of Chicago Press.

Shiffrin, R.M. (1996). Laboratory experimentation on the genesis of expertise. In K.A. Ericsson (Ed.), *The road to excellence* (pp. 337-345).

Shipley, W.C. (1940). A self-administering scale for measuring cognitive impairment and deterioration. *Journal of Psychology, 9,* 371-377.

Shirley, D.A., and Langan-Fox, J. (1996). Intuition: a review of the literature. *Psychological Review, 79,* 563-684.

Siegler, R.S., and Richards, D.D. (1982). The development of intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence.* New York: Cambridge University Press.

Simon, H.A., and Chase, W.G. (1973). Skill in chess. *American Scientist, 61,* 391-403.

Simon, H.A., and Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 8,* 165-190.

Simonton, D.K. (1980). Intuition process and analysis: a predictive and explanatory model. *Genetic Psychology Monographs, 102,* 3-60.

Simonton, K.A. (1996). Creative expertise: A life-span developmental perspective. In K.A. Ericsson (Ed.), *The road to excellence* (pp. 227-253).

Sinnott, J.D. (1989). A model for solution of ill-structured problems: implications for everyday and abstract problem solving. In J.D. Sinnott (Ed.), *Everyday problem solving: Theory and applications.* New York: Praeger.

Smith, J. and Baltes, P.B. (1990). Wisdom-related knowledge: Age/cohort differences in response to life-planning problems. *Developmental Psychology, 26,(3),* 494-505.

Smith, P.C., and Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology, 47,* 149-155.

Snow, R.E., and Lohman, D.F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology, 76,* 347-376.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30,* 526-537.

Spearman, C.E. (1904). 'General intelligence' objectively determined and measured. *American Journal of Psychology, 15,* 201-293.

Spearman, C. (1923). *The nature of 'intelligence' and the principles of cognition* (2nd ed.). London: Macmillan. (1923 edition reprinted in 1973 by Arno Press, New York).

Spearman, C. (1927). *The abilities of man.* London: Macmillan.

Sperry, R.W. (1961). Cerebral organization and behavior. *Science, 133,* 1749-1757.

Sternberg, R.J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities.* Hillsdale, NJ: Erlbaum.

Sternberg, R.J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology, 73,* 1-16.

Sternberg, R.J. (1983). Components of human intelligence. *Cognition, 15,* 1-48.

Sternberg, R.J. (1985a). *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge University Press.

Sternberg, R.J. (1985b). *Human abilities: An information-processing approach.* San Francisco: Freeman.

Sternberg, R.J. (1986). *Intelligence applied: Understanding and increasing your cognitive abilities.* San Diego: Harcourt Brace Jovanovich.

Sternberg, R.J. (1987). Most vocabulary is learned from context. In M.G. McKeown and M.E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1988). *The triarchic mind: A new theory of human intelligence.* New York: Viking.

Sternberg, R.J. (1990a). *Metaphors of mind: Conceptions of the nature of intelligence.* New York: Cambridge University Press.

Sternberg, R.J. (1991a). Theorybased testing of cognitive abilities: Rationale for the Triarchic Abilities Test. In H. Rowe (Ed.), *Cognition: Reconceptualization and measurement* (pp. 183-202). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1991b). Triarchic abilities test. In D. Dickinson (Ed.), *Creating the future: Perspectives on educational change* (pp. 76-81). Aston Clinton, Bucks, UK: Accelerated Learning Systems.

Sternberg, R.J. (1993). *Sternberg Triarchic Abilities Test.* Unpublished test.

Sternberg, R.J. (1994a). Cognitive conceptions of expertise. *International Journal of Expert Systems: Research and Applications*, *7* (1), 1-12.

Sternberg, R.J. (1994b). Intelligence. In R.J. Sternberg (Ed.), *Handbook of perception and cognition: Thinking and problem solving* (pp. 263-288). San Diego: Academic Press.

Sternberg, R.J. (1994c). Tacit knowledge and job success. In N. Anderson and P. Herriot (Eds.), *Assessment and selection in organizations* (pp. 27-39). New York: Wiley.

Sternberg, R.J. (1995a). Expertise in complex problem solving: A comparison of alternative conceptions. In P.A. Frensch and J. Funke (Eds.), *Complex problem solving: European perspectives* (pp. 295-321). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1995b). *In search of the human mind.* Orlando: Harcourt Brace College Publishers.

Sternberg, R.J. (1995c). Theory and measurement of tacit knowledge as part of practical intelligence. *Zeitschrift fur Psychologie, 203,* 319-333.

Sternberg, R.J. (1996). Costs of expertise. In K.A. Ericsson (Ed.), *The road to excellence.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1997a). *Successful intelligence.* New York: Plume.

Sternberg, R.J. (1997b). Tacit knowledge and job success. In N. Anderson and P. Herriot (Eds), *International handbook of selection and assessment* (pp. 201-213). New York: Wiley.

Sternberg, R.J. (1998a). Abilities as forms of developing expertise. *Educational Researcher, 27,* 11-20.

Sternberg, R.J. (1998c). Successful intelligence: An expanded approach to understanding intelligence. In K. Pribram (Ed.), *Brain and Values: Is a biological science of values possible* (pp.1-21). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J. (1999). A propulsion model of types of creative contributions. *Review of General Psychology, 3,* 83-100.

Sternberg, R.J. (in press-a). Intelligence as developing expertise. *Contemporary Educational Psychology*.

Sternberg, R.J. (in press-b). The theory of successful intelligence. *Review of General Psychology*.

Sternberg, R.J., and Ben-Zeev, T. (in press). *Complex cognition: The psychology of human thought*. New York: Oxford University Press.

Sternberg, R.J., and Berg, C.A. (1986). Quantitative integration: Definitions of intelligence: A comparison of the 1921 and 1986 symposia. In R.J. Sternberg and D.K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 155-162). Norwood, NJ: Ablex.

Sternberg, R.J., and Berg, C.A. (1992). Intellectual *development*. New York: Cambridge University Press.

Sternberg, R.J., and Clinkenbeard, P.R. (1995). A triarchic model of identifying, teaching, and assessing gifted children. *Roeper Review, 17 (4)*, 255-260.

Sternberg, R.J., Conway, B.E., Ketron, J.L., and Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41*, 37-55.

Sternberg, R.J., and Detterman D.K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition.* Norwood, NJ: Ablex.

Sternberg, R.J., Ferrari, M., Clinkenbeard, P.R., and Grigorenko, E.L. (1996). Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly, 40*(3), 129-137.

Sternberg, R.J., Forsythe, G.B., Hedlund, J., Horvath, J.A., Wagner, R.K., Williams, W.M., Snook, S., and Grigorenko, E.L. (in press). *Practical intelligence in everyday life.* New York: Cambridge University Press.

Sternberg, R.J., and Gardner, M.K. (1982). A componential interpretation of the general factor in human intelligence. In H.J. Eysenck (Ed.), *A model for intelligence* (pp. 231-254). Berlin: SpringerVerlag.

Sternberg, R.J., and Grigorenko, E.L. (1997a). The cognitive costs of physical and mental ill health: Applying the psychology of the developed world to the problems of the developing world. *Eye on Psi Chi, 2*, 20-27.

Sternberg, R.J., and Grigorenko, E.L. (Eds.). (1997b). *Intelligence, heredity and environment.* New York: Cambridge University Press.

Sternberg, R.J., Grigorenko, E.L., and Ferrari, M. (in press). Developing expertise: The respective roles of abilities and deliberate practice. In M. Ferrari (Ed.), *Expertise and its development*. New York: Guilford.

Sternberg, R.J., Grigorenko, E.L., Ferrari, M. and Clinkenbeard, P. (1999). A triarchic analysis of an aptitude-treatment interaction. *European Journal of Psychological Assessment*,*15*(1), 1-11.

Sternberg, R.J., Grigorenko, E.L., Ngorosho, D., Tantubuye, E., Mbise, A., Nokes, C., and Bundy, D.A. (1999). Hidden cognitive potential in rural Tanzanian school children. Unpublished manuscript.

Sternberg, R.J., and Horvath, J.A. (Eds.). (1999). *Tacit knowledge in professional practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R.J., and Kaufman, J.C. (1996). Innovation and intelligence testing: The curious case the dog that didn't bark. *European Journal of Psychological Assessment, 12*, 175-182.

Sternberg, R.J., and Kaufman, J.C. (1998). Human abilities. *Annual Review of Psychology, 49*, 479-502.

Sternberg, R.J., and Lubart, T.I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity.* New York: Free Press.

Sternberg, R.J., Nokes, K., Geissler, P.W., Prince, R., Okatcha, F., Bundy, D., and Grigorenko, E.L. (in press). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*.

Sternberg, R. J., Okagaki, L., and Jackson, A. (1990). Practical intelligence for success in school. *Educational Leadership, 48,* 35-39.

Sternberg, R.J., and Powell, J.S. (1983). Comprehending verbal comprehension. *American Psychologist, 38,* 878-893.

Sternberg, R.J., and Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition, 3,* 168-192.

Sternberg, R.J., and Wagner, R.K. (Eds.). (1986). *Practical intelligence: Nature and origins of competence in the everyday world.* New York: Cambridge University Press.

Sternberg, R.J., and Wagner, R.K. (1993). The g-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science, 2,* 1-5.

Sternberg, R.J., and Wagner, R.K. (Eds.). (1994). *Mind in context.* New York: Cambridge University Press.

Sternberg, R.J., Wagner, R.K., and Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese and J. Puckett (Eds.), *Advances in lifespan development* (pp. 205-227). Hillsdale, NJ: Erlbaum.

Sternberg, R.J., Wagner, R.K., Williams, W.M., and Horvath, J.A. (1995). Testing common sense. *American Psychologist, 50,* 912-927.

Strang, R. (1930). Measures of social intelligence. *American Journal of Sociology, 36,* 263-269.

Stricker, L.J., and Rock, D.A. (1990). Interpersonal competence, social intelligence, and general ability. *Personality and Individual Differences, 11,* 833-839.

Strough, J., Berg, C., and Sansone, C. (1996). Goals for solving everyday problems across the life span: age and gender differences in the salience of interpersonal concerns. *Developmental Psychology, 32,* 1106-1115.

Super, C.M., and Harness, S. (1982). The infants' niche in rural Kenya and metropolitan America. In L.L Adler (Ed.), *Cross-Cultural Research at Issue.* New York: Academic Press.

Terman, L.M. (1950). *Concept Mastery Test.* New York: The Psychological Corporation.

Thorndike, E.L. (1920). Intelligence and its use. *Harper's Magazine, 140,* 227-235.

Thorndike, R.L., and Stein, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychological Bulletin, 34,* 275-285.

Thornton, G.C., and Byham, W.C. (1982). *Assessment centers and managerial performance.* New York: Academic Press.

Thornton, G.C., and Cleveland, J.N. (1990). Developing managerial talent through simulation. *American Psychologist, 45,* 190-199.

Tornow, W.W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management, 32,* 221-229.

Vallacher, R.R., and Wegner, D.M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94,* 3-15.

Valsiner, J., and Leung, M.C. (1994). From intelligence to knowledge construction: a sociogenetic process approach. In R.J. Sternberg and R.K. Wagner (Eds.), *Mind in context* (pp. 202-217). New York: Cambridge University Press.

Vernon, P.E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology, 4,* 42-57.

Voss, J.F., Perkins, D.N., and Segal, J.W. (Eds.). (1991). *Informal reasoning and education.* Hillsdale, NJ: Erlbaum.

Vygotsky, L. (1978). *Mind in society: The development of higher order processes.* Cambridge, MA: Harvard University Press.

Wagner, D.A. (1978). Memories of Morocco: The influence of age, schooling and environment on memory. *Cognitive Psychology, 10,* 1-28.

Wagner, R.K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52,* 1236-1247.

Wagner, R.K. (1991). Managerial problem-solving. In R.J. Sternberg and P. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 159-183). Hillsdale, NJ: Erlbaum.

Wagner, R.K. (1997). Intelligence, training, and employment. *American Psychologist, 52,* 1059-1069.

Wagner, R.K. (in press). Practical intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence.* New York, NY: Cambridge University Press.

Wagner, R.K., Rashotte, C.A., and Sternberg, R.J. (1994). *Tacit Knowledge in sales: Rules of thumb for selling anything to anyone.* Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Wagner, R.K., and Sternberg, R.J. (1985). Practical intelligence in realworld pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49,* 436-458.

Wagner, R.K., and Sternberg, R.J. (1986). Tacit knowledge and intelligence in the everyday world. In R.J. Sternberg and R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51-83). New York: Cambridge University Press.

Wagner, R.K., and Sternberg, R.J. (1989). *Tacit Knowledge Inventory for Sales: Written.* Unpublished test.

Wagner, R.K., and Sternberg, R.J. (1990). Street smarts. In K.E. Clark and M.B. Clark (Eds.), *Measures of leadership* (pp.493-504). West Orange, NJ: Leadership Library of America.

Wagner, R.K., and Sternberg, R.J. (1991). *Tacit Knowledge Inventory for Managers.* San Antonio, TX: The Psychological Corporation.

Wagner, R.K., Sujan, H., Sujan, M., Rashotte, C.A. and Sternberg, R.J. (1999). Tacit knowledge in sales. In R.J. Sternberg and J.A. Horvath (Eds.), *Tacit knowledge in professional practice* (pp. 155-182). Mahwah, NJ: Lawrence Erlbaum Associates.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams and Wilkins.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York: Psychological Corporation.

Wedeck, J. (1947). The relationship between personality and psychological ability. *British Journal of Psychology,* 36, 133-151.

Wertsch, J.V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

Wertsch, J., and Kanner, B.G. (1994). A sociocultural approach to cognitive development. In R.J. Sternberg and C.A. Berg (Eds.), *Cognitive development* (pp. 328-349). New York, NY: Cambridge University Press.

Wigdor, A.K. and Garner, W.R. (Eds.). (1982). *Ability testing: Uses, consequences, and controversies.* Washington, DC: National Academy Press.

Williams, S.A., Denney, N.W., and Schadler, M. (1983). Elderly adults' perception of their own cognitive development during the adult years. *International Journal of Aging and Human Development, 16,* 147-158.

Williams, W.M., Blythe, T., White, N., Li, J., Sternberg, R.J., and Gardner, H.I. (1996). *Practical intelligence for school: A handbook for teachers of grades 5-8*. New York: HarperCollins.

Willis, S.L. (1987). Cognitive training and everyday competence. In K.W. Schaie (Ed.), *Annual Review of Gerontology and Geriatrics* (Vol. 7, pp. 159-188). New York, NY: Springer.

Willis, S.L. (1989). Improvement with cognitive training: Which dogs learn what tricks? In L.W. Poon, D.C. Rubin, and B.A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 300-329). New York, NY: Cambridge University Press.

Willis, S.L. (1990). Contributions of cognitive training research to late life potential. In M. Perlmutter (Ed.), *Late life potential* (pp. 25-42). Washington, DC: Gerontological Society of America.

Willis, S.L., Blieszner, R., and Baltes, P.B. (1981). Training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology, 73,* 41-50.

Willis, S.L., and Nesselroade, C.S. (1990). Long-term effects of fluid ability training in old-old age. *Developmental Psychology, 26,* 905-910.

Willis, S.L., and Schaie, K.W. (1986). Practical intelligence in later adulthood. In R.J. Sternberg and R. Wagner (Eds.), *Practical intelligence* (pp. 236-270). New York: Cambridge University Press.

Willis, S.L. and Schaie, K.W. (1994). Cognitive training in the normal elderly. In F. Boller (Ed.), *Cerebral plasticity and cognitive stimulation.* New York: Springer-Verlag.

Wong, C.T., Day, J.D., Maxwell, S.E., and Meara, N.M. (1995). A multitrait-multimethod study of academic and social intelligence in college students. *Journal of Educational Psychology, 87,* 117-133.

Worden, P.E., and Sherman-Brown, S. (1983). A word-frequency cohort effect in young versus elderly adults-memory for words. *Developmental Psychology, 19,* 521-530.

Yang, S., and Sternberg, R.J. (1997). Conceptions of intelligence in ancient Chinese philosophy. *Journal of Theoretical and Philosophical Psychology, 17(2),* 101-119.

Yerkes, R.M. (1921). *The measurement and appraisal of adult intelligence (4th ed.)*. Baltimore, MD: Williams and Wilkins.

Yussen, S.R., and Kane, P. (1985). Children's concept of intelligence. In S.R. Yussen (Ed.), *The growth of reflection in children.* New York: Academic Press.

Chapter 9

The ICT Literacy Framework

Irwin Kirsch *Educational Testing Service Princeton, NJ*
Marilou Lennon *Educational Testing Service Princeton, NJ*

Table of contents

1. Introduction

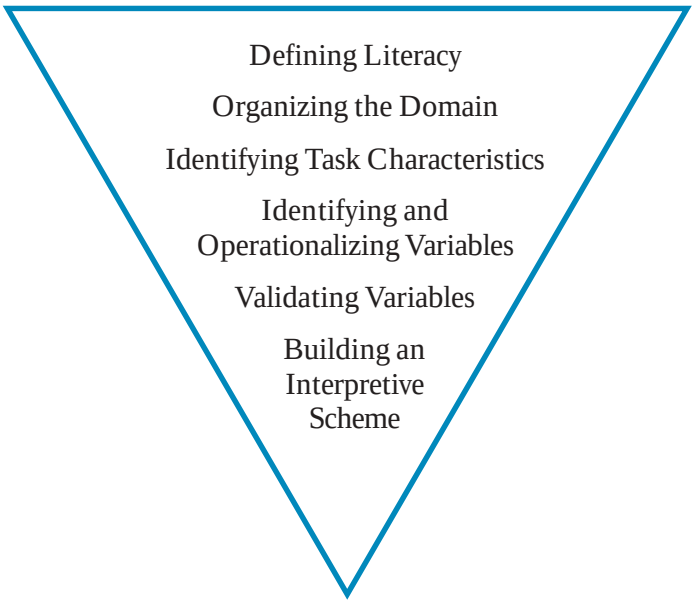
What does it mean to be a literate member of society? The growing acceptance of lifelong learning has expanded the views and demands of literacy. Literacy is no longer seen as a condition that one either has or is lacking. Rather, it is seen as a continuum of knowledge, skills, and strategies that individuals acquire over the course of their lives in various contexts and through interactions with their peers and with the larger communities in which they participate. As historians remind us, literacy in its earliest form consisted of little more than being able to sign one's name on a legal document. It was not until later that fluent oral reading became important and not until the 20th century that reading to gain information was given primary emphasis. As we move into the 21st century, our conception of literacy is evolving once again. The prevalence of technology in the everyday lives of the world's citizens has grown at a rate that many would have found hard to imagine 25 or even 10 years ago. Policy makers, business leaders, and educators have come to expand their notion of a literate populace to include the skills and abilities that will enable citizens to function in an increasingly technological world.

2. Developing a framework

The task of the International ICT Literacy Panel was to develop a framework that would define ICT literacy and provide the foundation for the design and conduct of large-scale assessments and diagnostic tests. While the chief benefit of developing a framework for ICT literacy is improved measurement, a number of other potential benefits are also seen as important. Namely,

- A framework provides a common language and a vehicle for discussing the definition and assumptions surrounding the domain.
- Such a discussion provides a mechanism for building consensus around the framework and measurement goals that grow from it.
- We construct a better understanding of what is being measured through the process of developing the framework and linking it to evidence collected from assessment tasks.
- This understanding and its connection to what we say about learners provides an important link between public policy, assessment, and research which furthers the utility of the data that are collected.

To accomplish this task, the panel chose to adopt the process used to develop frameworks for the International Adult Literacy Survey (OECD and Statistics Canada 1995; OECD and Development and Human Resources Development Canada 1997; OECD and Statistics Canada 2000) and for the Reading Literacy Survey conducted as part of PISA, the Programme for International Student Assessment (OECD, 1999). This process consists of six steps, shown in the following diagram and explained more fully below (Kirsch 2001).



1. The first step is to develop a working definition of the domain including the assumptions underlying it. Before the definition is developed, the domain and the skills and abilities it encompasses are wide open. It is the definition that sets the boundaries for what will be measured and what will not.

2. Once the definition is developed, it is important to think about the kinds of tasks that represent the skills and abilities included under that definition. Those tasks must then be categorized, or organized, to inform test design and result in meaningful score reporting. Step 2 allows one to move beyond a laundry list of tasks or skills to a coherent representation of the domain that will permit policy makers and others to summarize and report information in more useful ways.
3. Step 3 involves identifying a set of key characteristics that will be used in constructing tasks for the assessment. This may include characteristics of the stimulus materials to be used as well as characteristics of the tasks presented to examinees.
4. In step 4, the variables associated with each task characteristic are specified.
5. In step 5, research is conducted to show which variables account for large percentages of the variance in the distribution of tasks and thereby contribute most towards understanding task difficulty and predicting performance.
6. Finally in step 6, an interpretative scheme is built that uses the validated variables to explain task difficulty and examinee performance. The work of this panel involved the first two steps: defining ICT literacy and organizing the domain.

3. Defining ICT literacy

The International ICT Literacy Panel was comprised of educators, technology experts, scholars and industry and labor representatives from Australia, Brazil, Canada, France, and the United States. Our deliberations resulted in the following definition:

> ICT literacy is using digital technology, communications tools, and/or networks to access, manage, integrate, evaluate and create information in order to function in a knowledge society.

This definition carries several assumptions made by the panel and therefore it is important to consider each part of the definition in turn.

"ICT..."

Information Technology (IT) has been used for many years, particularly in the United States, and refers to the electronic display, processing, and storage of information, but not necessarily the transmission of the information. The term carries strong historical associations with enterprise data processing and centralized computer services.

However, Information and Communication Technology (ICT) represents the set of activities and technologies that fall into the union of IT and communication technologies. Global industry, international media, and academics increasingly now use ICT to describe this union. The real benefit of adding "communication" doesn't derive from including specific technologies, such as routers or servers, but from the dynamism implicit in interconnected social, economic, and information networks. ICT is characterized by unprecedented global flows in information, products, people, capital, and ideas. These flows are enabled by ICT: their sheer scale and pace would not be possible without the ability to connect vast networks of individuals across geographic boundaries at negligible marginal cost.

"...literacy is..."

The panel selected the term literacy over other terms such as competency, ability, or fluency that have been used in earlier frameworks (Committee on Information Technology Literacy, 1999). To some "literacy" connotes functional literacy and implies basic or fundamental skills. To the panel, the term literacy implies a universal need, a condition that must be met to enable full and equitable economic and social participation. We view literacy as a tool that may be applied to simple or more complicated contexts — like a hammer that can be used to build a shelf, or a house. In its broadest sense, literacy is a dynamic tool that allows individuals to continuously learn and grow.

The increasing role of technology in our lives requires us to expand our notion of literacy. It is obvious that to function fully and effectively in society, individuals must be literate in terms of traditional domains such as reading and numeracy. But today it is becoming increasingly clear that ICT literacy joins the ranks of essential and fundamental requirements. Perhaps as important is the panel's belief that those who fail to acquire this new kind of literacy, like the more traditional literacy skills, will find themselves falling further behind as economies and societies grow and change over the years ahead.

"...using digital technology, communications tools, and/or networks..."

The description of digital technology, communication tools, and/or networks reflects the same thinking that stimulated the panel's use of information and communication technology (ICT) versus information technology (IT). Digital technology reflects

hardware and software products, communication tools reflect those products and services used to transmit information, and networks themselves are the pathways for this transmission. The words are meant to be as inclusive as possible to reflect the breadth of hardware, software, and infrastructures that makeup ICT.

"...to access, manage, integrate, evaluate and create information..."

Technology is used for an ever-increasing range of purposes to accomplish many different kinds of tasks. This phrase is meant to reflect that range as well as to define five critical components of ICT literacy. The five components represent a continuum of skills and knowledge and are presented in a sequence suggesting increasing cognitive complexity. After discussions regarding the kinds of tasks represented by each component, the panel agreed to the following definitions:

- Access—knowing about and knowing how to collect and/or retrieve information.
- Manage—applying an existing organizational or classification scheme.
- Integrate—interpreting and representing information. It involves summarizing, comparing and contrasting.
- Evaluate—making judgments about the quality, relevance, usefulness, or efficiency of information.
- Create—generating information by adapting, applying, designing, inventing, or authoring information.

"...in order to function in a knowledge society."

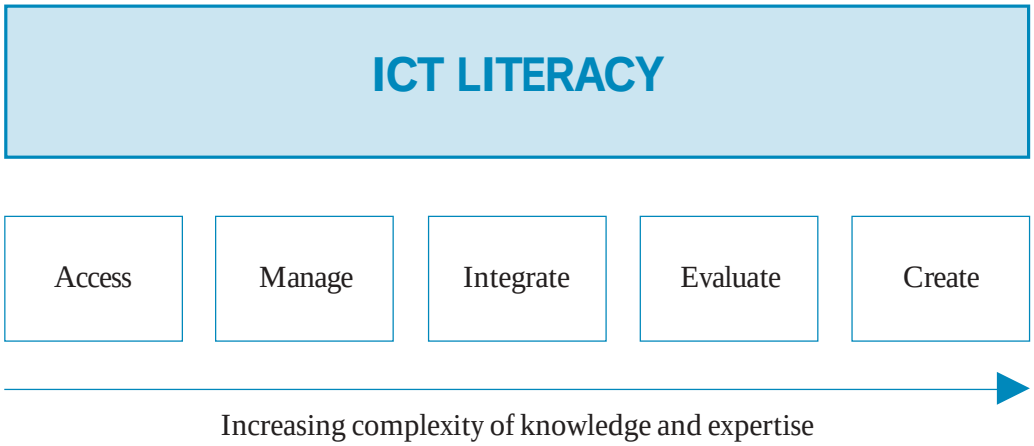
This phrase reflects the range of contexts in which individuals will be able to apply their ICT literacy—from defined ones such as graduating from school or functioning on a job to those which are less defined and less concrete but which can extend and enrich one's personal life. The phrase "in order to function" is meant to acknowledge the fact that ICT literacy will provide individuals with a means of contributing to and benefiting from economically developed or developing societies. We believe that ICT literacy skills are becoming increasingly important not only for nations to maintain or improve their standard of living but for the well being of individuals as well. The phrase "in a knowledge society" refers to the changing nature of cultures in the 21st century—an age in which ideas and information are increasingly the drivers of progress. The expanding roles of technology and access to information on a global scale have the potential to change, and hopefully improve, the way we live, learn and work.

4. Organizing the domain

Once we had defined what was meant by ICT literacy and laid out the assumptions underlying that definition, the next step was to develop an organizing framework for ICT literacy. This is an important step because the way in which the domain is organized affects test design and the kinds of tasks that will be developed to provide evidence about the status of ICT literacy in a population of interest. The panel's task was to define the critical organizing categories for the domain of ICT literacy and how they were related.

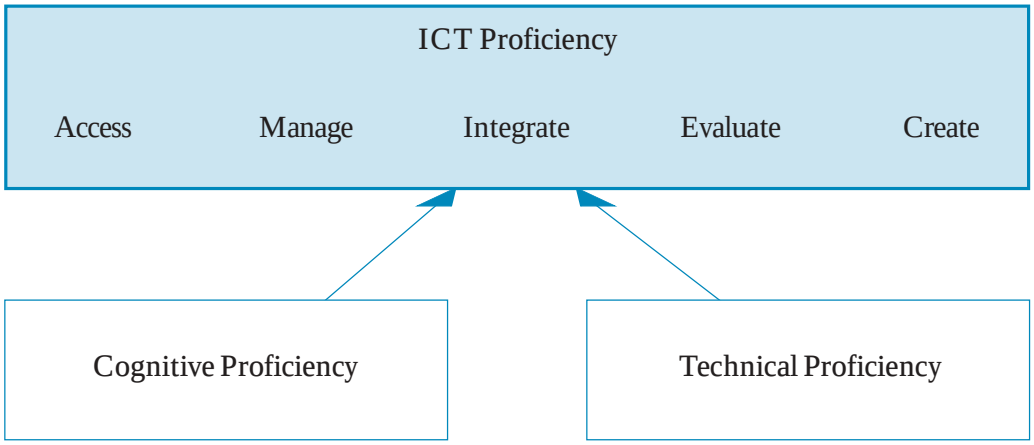
In our definition of ICT literacy, we identified five components we view as essential for functioning in a knowledge society: accessing, managing, integrating, evaluating and creating information in a technology context. These components, represented in Figure 1, formed the initial organizational scheme for the domain of tasks that make up ICT literacy.

Figure 1



Upon further consideration, however, we chose to expand this unidimensional model to more fully represent the complexity of ICT literacy. This organizational scheme, shown below in Figure 2, illustrates the foundational set of skills and knowledge that underlie ICT literacy: cognitive and technical proficiency.

Figure 2



The three proficiencies are defined as follows.

- Cognitive Proficiency—the desired *foundational skills* of everyday life at school, at home, and at work. Literacy, numeracy, problem solving, and spatial/visual literacy demonstrate these proficiencies.
- Technical Proficiency—the basic components of digital literacy. It includes a foundational knowledge of hardware, software applications, networks, and elements of digital technology.
- ICT Proficiency—the integration and application of cognitive and technical skills. ICT proficiencies are seen as enablers; that is, they allow individuals to maximize the capabilities of technology. At the highest level, ICT proficiencies result in innovation, individual transformation, and societal change.

As conceived in this framework, ICT literacy includes both cognitive and technical proficiency. For example, in order to successfully perform an ICT task such as searching the Internet to find and compare treatment options for a medical condition, an individual must apply reading and problem solving skills (cognitive) and be able to access information on the Internet using a search engine (technical). While cognitive and technical proficiencies are both necessary components of ICT literacy, each is a distinct domain. Cognitive and technical proficiency each represent independent domains in which the associated knowledge and skills interact to influence ICT literacy. An example is shown below in Figure 3.

Figure 3

	Low technical proficiency	High technical proficiency
High Cognitive Proficiency	A	B
Low Cognitive Proficiency	C	D

One would expect that individuals with low cognitive proficiency but high technical proficiency (cell 'D') would be able to perform particular technical tasks in which they had been trained. However, they would probably not possess the kind of generalizable skills or knowledge that could help them work with new applications or perform novel tasks and they would most likely not be able to acquire such skills independently. People with high cognitive proficiency but low technical proficiency (cell 'A') would require technical training (and possibly motivation or practice time) in order to develop ICT proficiency but would be expected to do so and once engaged with ICT would be able to acquire new skills and knowledge quickly and independently.

The representation of ICT literacy shown in Figure 2 provides an organizational scheme for both developing large-scale assessments or individual tests and evaluating existing measures. The framework leaves open the possibility that different constituencies could develop different assessments or individual tests for different purposes by focusing on various parts of the framework itself and by defining the kinds of evidence that might be associated with each. For the purposes of discussion, we present three types of assessments that might be developed using the ICT literacy framework: holistic, component and diagnostic. Each is described in turn below. Additional detail about possible assessment tasks can be found in Appendices A, B and C. These sample tasks are also available in a more interactive form at www.ets.org/research/ictliteracy/index.html.

A holistic assessment would be of most interest to constituencies who wished to focus on how well a test taker completed a given set of tasks rather than on the component skills that make up those tasks. An example would be a task that required test takers to

create a flyer for a neighborhood clean-up day. Specific task requirements as well as information such as when and where the event was to be held would be provided and test takers would be scored on how well they completed the final product. (A more detailed illustration of this task can be found in Appendix A.) Such an assessment would allow one to rank order groups or individuals and to make general statements about their ICT knowledge and skills.

Alternatively, one might choose to build an assessment that independently measured the knowledge and skills associated with each of the five components of ICT Proficiency (access, manage, evaluate, integrate and create). This component assessment would result in a measure that could provide general information about the kinds of generative ICT tasks a population, or an individual, could perform. An example of one such task and its associated proficiencies is presented below. (For a more detailed illustration of this task, see Appendix B.)

Scenario: Employees in your company have been asked to provide information about the technology training courses they have taken during the past year. They have sent e-mail messages to their supervisors and each supervisor has forwarded the information to the director of human resources. You've been asked to organize the information, evaluate the extent to which company-based courses are being utilized, and make a recommendation to the human resources department about which courses should be continued next year.

Access	Select and open appropriate e-mails from inbox list.
Manage	Identify and organzie the relevant information in each e-mails.
Integrate	Summarize the interest in the courses provided by the company.
Evaluate	Decide which courses should be continued next year, based on last year's attendance.
Create	Write up your recommendation in the form of an e-mails to the vice president of human resources.

But if an individual or a group of individuals performed poorly on this measure, one would be hard pressed to understand or explain why. Were there underlying reading or language problems? Did test takers have sufficient technical knowledge to complete the tasks presented in the ICT measures? To understand what role these other domains contributed one would have to include cognitive and technical tasks in the assessment or test. Alternatively, one might want to focus on particular ICT proficiencies (for example, how well a person can access and manage information) and their underlying cognitive and technical components. This would involve creating tasks that measured these types of skills and knowledge across the three proficiency domains. These measures would provide evidence separating literacy and technology proficiencies from ICT proficiency. Such information would be useful for constituencies such as adult basic education centers interested in diagnosing and remediating problems students are having accessing information on the Internet. A series of tasks that might be appropriate in this context are presented below (and in more detail in Appendix C).

Scenario: Following a stroke, your mother has been diagnosed with an atrial septal defect, or a hole in one section of her heart. While not an emergency, her doctor has recommended open-heart surgery to repair the hole and reduce the risk of additional strokes. You would like to find several reliable sources on the Web that recommend treatment options for this condition.

Access	Using a search engine, locate sites that have articles about holes in the heart, or atrial septal defects.

Students having trouble with this basic ICT task could be presented with related cognitive and technical tasks to help diagnose what was causing their difficulty. For example, students might be presented with multiple-choice questions asking them to select the best word or phrase to use when searching for some specified information. Included among the choices might be terms that are overly general or specific. Students having difficulty with this type of task might need practice in defining categories and efficient search strategies. In addition, very basic computer tasks, such as opening a search engine, clicking on sites, and navigating back to the search engine from those sites, might uncover technical skills requiring review or training.

Currently, there are various measures of literacy, numeracy and problem solving being used in large-scale assessments of school age and adult populations. There is also a measure of technical knowledge and under-standing that is being used with school-aged populations. These are traditional paper and pencil measures. No attempt has been made, however, to build computer-based tasks to measure the integration of these cognitive and technical domains or to separate out the role each plays in the development of these more generative ICT proficiencies. The panel believes that the measurement of ICT literacy using paper and pencil will limit the ability to assess the full domain of knowledge and skills. Valuable information will be lost if assessment tasks are not embedded in real-world settings incorporating technology. For example, the measurement of an individual's ability to search for and access information would be hindered if the measurement did not provide an opportunity to log onto the Internet or a similar type of environment.

5. Next steps

As the panel began its deliberations about ICT literacy and how should it be defined and operationalized, we soon recognized that many of our discussions focused around the issue of the digital divide. This divide is commonly defined in terms of connectivity and the inequalities of access that exist both within and across countries. The more important issue the panel recognized was that the true potential of ICT—that is, the ability to transform individuals and societies—came not just from being wired together but also from having the knowledge and skills to use technology and to understand the roles it can play in our lives. As the president of Morris Brown College recently stated, "Merely having access to a box—an information box—does not necessarily mean that you have improved, or that you're more literate, or that you're better able to solve problems in the community" (Young, 2001).

This perspective led the panel to determine what they saw as the important issue facing us as society continues to invest in technologies and as technology continues to alter the way we work and live our lives. Then we wanted to use this storyline as a lead in to the definition of ICT literacy and how it should be operationalized into a framework. This report has taken the initial steps in building a framework by providing a consensus definition of ICT literacy and a model that can be used to further operationalize this construct for a variety of purposes.

The next steps will involve defining the kinds of evidence that should be gathered with respect to each level of the model—ICT, cognitive and technical proficiencies—and the kinds of activities that would elicit that evidence. This evidence and related activities will vary depending on the purpose of the planned assessment or test.

The framework begun with this paper, along with a prototype of online tasks, will allow ETS to discuss the potential for large-scale assessments or individualized tests with potential sponsors. The major stakeholders who will be interested in this framework and its resulting assessments are international and diverse, and therefore create a unique challenge as well as opportunity. They include government policy makers, corporate leaders, industry associations, unions, workforce groups, educators (K-12, higher education, national educational associations, researchers), consumer and public interest groups, and relevant international associations. The buy-in, cooperation, and support of these groups will be essential in the achievement of global ICT literacy.

ICT has become a permanent part of everyday life. It fundamentally changes how we live, learn, and work. Because ICT is considered an emerging and fundamental literacy, significant attention must be paid to insuring that all citizens have access and opportunity to gain the needed skills to function effectively in a knowledge society.

Appendix A

Sample assessment task—ICT proficiency

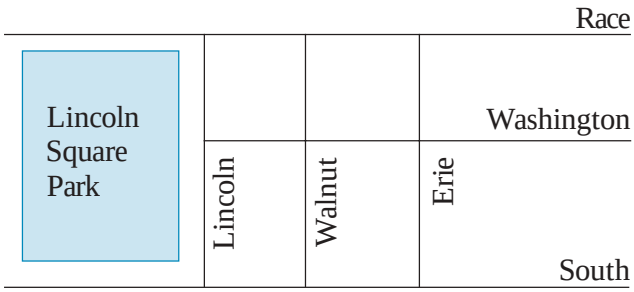
- Holistic assessment of ICT skills and knowledge
- Scenario presented along with a variety of tools (spreadsheet, word processor, etc.)

In this type of assessment, test takers would be evaluated solely on the end product they created (for example, a database, presentation, or document). Component skills would not be isolated and individually assessed. Instead, a scoring scheme would be developed which defined levels of performance and the criteria for reaching each level. This scheme would represent the collective judgments of experts in the field about what adults should know and be able to do in the ICT domain. Below is an example of what one task in a holistic assessment might look like. A complete assessment would include a number of different tasks that vary in difficulty and require a range of ICT knowledge and skills.

Opening scenario (Community context)

You've volunteered to create a flyer for a community clean-up day to be held in your neighborhood. Include the map below along with the following information and create an attractive one-page flyer for the event.

The event will take place on Saturday, May 6th from 1:00 until 4:00. Volunteers are being asked to meet at Lincoln Square Park. Event organizers would like a tear-off registration slip to be included on the flyer where volunteers



can print their name, address and phone number. The registration forms should be dropped off at the community center on Race Street by May 1st.

To complete this task, test takers would need to use a word processing program to create a flyer. The final product would be scored on the accuracy and completeness of the information it contained (e.g., did the flyer include all the relevant information about dates and times, the map and the tear-off registration form?). Additional scoring points might include evaluating the layout and inclusion of graphic elements (borders, lines, etc.).

Appendix B

Sample assessment task—ICT components

- Focus on the components of ICT proficiency: access, manage, integrate, evaluate, and create
- Present a scenario followed by tasks addressing each of the five components

Below is an example of what one task in this type of assessment might look like. A complete assessment would include a number of different tasks that vary in difficulty and require a range of knowledge and skills in technical, cognitive, and problem-solving domains.

Opening scenario (Workplace context)

Employees in your company have been asked to provide information about the technology training courses that they have taken during the past year. They have sent e-mail messages to their supervisors and each supervisor has forwarded the information to the director of human resources. You've been asked to organize the information, evaluate the extent to which company-based courses are being utilized and make a recommendation to the human resources department about which courses should be continued next year.

The five components

Based on this scenario, test takers would be presented with a series of tasks. Each task would be designed to measure one of the five components, as summarized in the chart below. While the sequence in which individual test takers undertake these tasks might vary, each component could be scored discretely in order to better understand its relative contribution to an individual's overall ICT proficiency.

Access	Select and open appropriate e-mails from inbox list.
Manage	Identify and organzie the relevant information in each e-mails.
Integrate	Summarize the interest in the courses provided by the company.
Evaluate	Decide which courses should be continued next year, based on last year's attendance.
Create	Write up your recommendation in the form of an e-mails to the vice president of human resources.

Test takers might work from a screen that presents all of the task components and allows them to select the order in which they complete those tasks. An alternate approach would be have test takers work through a structured series of tasks with the first component presented, followed by the second component and so on.

A more detailed description of the component tasks is presented below.

Access and manage task

> Task Description: *Seven supervisors have sent information about training courses to Ann Simpson, Director of Human Resources, and she has forwarded them to you. Find and open each of those e-mails in your inbox. Select the text from each e-mail that provides information about training course attendance and copy it all into a single file.*

Test takers would be presented with a simulated inbox, similar to the sample shown below. Some might chose to open all the e-mails and then select the relevant information. Others might open one e-mail, select the critical information and then move on to the next. Whatever the sequence, to complete the task correctly test takers would be expected to open each of the correct e-mail messages and paste all the relevant information into a file.

Sample In-Box:

From	Subject	Received	Size
Simpson, Ann	FW: Training	12/17/01 10:32 AM	3 KB
Simpson, Ann	FW: Course Information	12/17/01 10:44 AM	2 KB
Davidson, Denise	RE: Lunch	12/17/01 10:57 AM	7 KB
Simpson, Ann	Work Objectives	12/17/01 11:11 AM	5 KB
Simpson, Ann	FW: Classes Taken	12/17/01 11:27 AM	3 KB
Corporate Communique	Virus Alert	12/17/01 12:01 PM	4 KB
Simpson, Ann	FW: Courses This Year	12/17/01 12:15 PM	4 KB
Simpson, Ann	FW: Training Classes	12/17/01 12:49 PM	2 KB
Gonzalez, Frank	Team meeting	12/17/01 1:08 PM	8 KB
Simpson, Ann	FW: Thursday Staff Meeting	12/17/01 1:11 PM	3 KB
Simpson, Ann	FW: Training Courses	12/17/01 1:59 PM	2 KB
Salverston, Amy	RE: Phone Billing	12/17/01 2:14 PM	6 KB
Mirano, Leslie	Training Class Question	12/17/01 2:48 PM	5 KB
Jenkins, Ralph	Update	12/17/01 3:19 PM	3 KB
Simpson, Ann	Memo for Davidson	12/17/01 3:21 PM	4 KB
Ellis, Edward	Re: Phone Conference	12/17/01 3:56 PM	2 KB
Simpson, Ann	FW: Staff Training Courses	12/17/01 4:17 PM	2 KB
Rogers, Charlie	FW: Memo Format	12/17/01 4:45 PM	3 KB

Ann – Jason and I met yesterday and have a schedule for the next team meetings. We will send that information out to everyone later today. Here is the information you requested about training courses. In my area, 25 people took one or more training classes this year. 15 people took Learning Excel, Level 1 (March 27 and 28), 20 took Introduction to Outlook (June 3 and 4) and 5 took Flash, Level 2 (October 19 and 20). The first two courses were given on site and the last was at the community college. We have gotten particularly positive feedback about the Outlook course. Let me know if you need any additional information.

- E. O'Brien

Integrate

Task Description: *You want to look at all the information the supervisors have provided so that you can see which of the courses taught at the company were most popular. Represent that information in a way that will help you make the recommendation about which courses to continue next year.*

Test takers would need to decide the best way to integrate and compare the information they have selected in the previous task. They might present the information in a list or series of lists, in a table, etc. In the sample response shown below, the information from the seven e-mail messages has been used to create a table that allows one to quickly compare course location and attendance across courses.

Sample response:

Location	Name of course	Number of employees
On site	Learning Excel	31
On site	Introduction to Outlook	50
On site	Visual Basic	5
On site	HTML	25
On site	Networking Essentials	2
Advantex Computer Training	C++	5
Community college	Flash, Level 2	5
Community college	Windows NT	17

Evaluate and create

Task Description: *Using last year's attendance figures for courses offered by the company, decide which courses should be offered next year. Write an e-mail to Ann Simpson with your recommendation, including as attachments any tables or charts that support your position.*

Test takers would need to identify the on-site courses with the best attendance based on the supervisor's reports. They would then to write up their recommendation and attach supporting documentation. Scoring models would be created to focus on the skills and knowledge deemed most relevant to assess for a particular population. For example, one might be interested in the extent to which test takers were able to support their recommendation with evidence from the original supervisor's e-mails, the sophistication of supporting documentation, or the test taker's ability to use software to create tables or graphs.

Appendix C

Sample assessment task—Diagnostic assessment

This type of assessment would allow one to investigate the cognitive and technical proficiencies underlying particular ICT components. On the surface, the Diagnostic Assessment would look exactly like the assessment of ICT Components. Only if and when test takers had difficulty with a component task would they see new types of tasks designed to assess underlying cognitive and technical skills.

The results of this kind of assessment could be used in a variety of ways:

- The assessment could provide an overall score of a person's ICT, cognitive and technical proficiency.
- A more detailed score reporting system might be developed that profiled specific strengths and weaknesses that an individual demonstrated.
- Links to existing or specially developed instructional materials could be provided to help teachers in education or training settings.
- Based on a person's performance, targeted instructional goals and suggestions on how best to reach those goals could be made available.

Just like in the ICT Component assessment, a number of scenarios in different contexts would be presented. One sample scenario, developed in a health context, is shown below.

Opening scenario (Health context)

Following a stroke, your mother has been diagnosed with an atrial septal defect, or a hole in one section of her heart. While not an emergency, her doctor has recommended open-heart surgery to repair the hole and reduce the risk of additional strokes. You would like to find several reliable sources on the Web that recommend treatment options for this condition.

The five components

Based on this scenario, test takers would be presented with a series of tasks organized around the five components, as summarized in the chart on the following page.

Access	Using a search engine, locate sites that have articles about holes in the heart, or atrial septal defects.
Manage	Evaluate the sites and identify three that you would expect to provide reliable medical information.
Integrate	Identify the treatment information in each article.
Evaluate	Compare and contrast the treatment options suggested in the articles.
Create	Devlop a Word document with treatments listed (citing sources) to share with physician.

As each task was completed it would be automatically scored. If a test taker did not complete a task correctly, related cognitive and technical tasks would be presented to try and determine if one or both of those areas were contributing to the individual's difficulty.

An example of how the Access task might be broken down follows.

Access task

Task Description: *Use the search engine provided to find three sites with information about your mother's medical condition as described in the opening scenario.*

Assessing underlying technical skills

If a test taker did not complete the task correctly, one question would be whether he or she had the requisite technical skills. Technically, this access task requires test takers to open up a browser, type a word or phrase into the text entry box, and click on the Search button. They might additionally need to open a site and then navigate back to search engine. As an individual test taker completed this task, the computer would record clicks, typing and other actions. Based on the test taker's responses, additional discrete computer-based tasks might be presented (e.g., "Type the phrase 'Movie Listings' into the search box" or "Click on the button that will take you back to the search page") to assess the technical skills underlying this basic Access task.

Assessing underlying cognitive skills

Cognitively, this access task requires a test taker to select or invent a search term that would yield the requested information. Some of the words or phrases in the task description and scenario would be more likely than others to provide the information needed. For example, typing in the phrase "hole in the heart" in one browser would yield the results shown below, none of which would be likely to include the information needed.

Typing in the more general term, "heart," would result in the following types of sites.

The word or phrase test takers used for their query would provide the basis for scoring this task (with a more precise phrase such as “atrial septal defect treatment” resulting in a higher score than “treating heart”). The program might also track if test takers refined their search based on the results each search yielded. If a test taker did not perform well on this task, other less open-ended tasks might be presented. These might include multiple-choice questions that asked test takers to select from the choices provided the best term to search for specified information. Questions that focused on general versus specific categories might provide additional diagnostic information. Another possibility would be to present a similar search task that was not computer based, such as locating specified information in the Yellow Pages and seeing if the test taker could successfully complete that task. The goal of any of these or additional follow-up tasks would be to try and identify underlying areas of difficulty that might be contributing to poor performance on the computer search task.

References

American Society of Training and Development (ASTD) *State of the industry report.* (2000). Alexandria, Virginia. Author.

Benton Foundation. (2000). *The e-rate in America: A tale of four cities*. Washington, DC: Author.

Blau, Andrew. (2001). *More than bit players: How information technology will change the way nonprofits and foundations work and thrive in the information age*. New York: Surdna Foundation.

Bollier, D. (2000). *Ecologies of innovation: The role of information and communications technologies: A report of the eighth annual Aspen Institute Roundtable on Information Technology*. Washington, DC: The Aspen Institute.

Cline, W. (1997). *Trade and income distribution*. Washington, DC: Institute for International Economics.

Commission on Technology and Adult Learning. (2001). *A vision of e-learning for America's workforce*. Alexandria, VA, and Washington, DC: American Society for Training and Development and the National Governors Association.

Committee on Information Technology Literacy. (1999). *Being fluent with information technology*. Washington, DC: National Academy Press.

de los Santos, G.E., de los Santos Jr., A.G., and Milliron, M.D. (Eds.). (2001). *Access in the information age: Community college bridging the digital divide*. Mission Viejo, CA: League for Innovation in the Community College.

Digital Opportunity Initiative. (2001). *Creating a development dynamic*. New York, NY: Accenture, Markle Foundation, United Nations Development Programme. http://www.opt-init.org/framework.html.

Educational Development Center. (2000). *IT pathway pipeline model: Rethinking information technology learning in schools*. Newton, MA: Author.

Ellis, Carlene. (2001). *Innovation in education: The increasing digital world—issues of today and tomorrow*. Presented at the National IT Workforce Convocation of the Information Technology Association of America, San Diego, CA. http://www.intel.com/education/community/contribution.

Goslee, S. (1998). *Losing ground bit by bit: Low-income communities in the information age*. Washington, DC: Benton Foundation.

Hart, D. (2001). Bridging the digital divide, *Hemispheres*.

Information Technology Association of America. (2001). *When can you start? Building better information technology skills and careers.* Arlington, VA: Author.

Information Technology Association of America. (2000). *Bridging the gap: Information technology skills for a new millennium*. Arlington, VA: Author.

Katz, S.N. (2001, June 15). In information technology, don't mistake a tool for a goal. *The Chronicle of Higher Education, The Chronicle Review, 47*(40), B7-B9. http://www.chronicle.com/ weekly/v47/i40/ 40b00701.htm.

Kirsch, I. (2001). *The International Adult Literacy Survey (IALS): Understanding what was measured* (ETS RR-01-25). Princeton, NJ: Educational Testing Service.

Morino Institute. (2001). *From access to outcomes: Raising the aspirations for technology initiatives in low-income communities*. Reston, Virginia: Author.

National Association of State Boards of Education. (2001). *Any time, any place, any path, any pace: Taking the lead on e-learning policy*. Alexandria, VA: Author.

National Telecommunications and Information Administration. (2000). *Falling through the net: Toward digital inclusion*. Washington, DC: The US Department of Commerce. Author.

National Telecommunications and Information Administration. (1999). *Falling through the net: Defining the digital divide*. Washington, DC: The US Department of Commerce. Author.

National Telecommunications and Information Administration. (1998). *Falling through the net II: New data on the digital divide*. Washington, DC: The US Department of Commerce. Author.

National Telecommunications and Information Administration. (1995). *Falling through the net: A survey of the "have nots" in rural and urban America*. Washington, DC: The US Department of Commerce. Author.

Organisation for Economic Co-operation and Development. (2001a). *Understanding the digital divide*. Paris: Author.

Organisation for Economic Co-operation and Development. (2001b). *The well-being of nations*. Paris: Author.

Organisation for Economic Co-operation and Development. (2000b). *Schooling for tomorrow: Learning to bridge the digital divide*. Paris: Author.

Organisation for Economic Co-operation and Development. (1999). *The economic and social impact of electronic commerce: Preliminary findings and research agenda*. Paris: Author.

Organisation for Economic Co-operation and Development and Human Resources Development Canada (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey.* Paris: Author.

Organisation for Economic Co-operation and Development and Statistics Canada (1995). *Literacy, economy and society: Results of the first International Adult Literacy Survey.* Paris and Ottawa: Author.

Organisation for Economic Co-operation and Development and Statistics Canada (2000). *Literacy in the information age: Final report of the International Adult Literacy Survey.* Paris and Ottawa: Author.

Reinhardt, A. (March, 1995). New ways to learn. *Byte Magazine*. http://www.byte.com/art/9503/sec7/ art1.htm.

Venezky, R.L. (2001). *Assessment of ICT concepts and skills: Summary and recommendations for PISA*. Unpublished manuscript, University of Delaware.

Young, J. (2001, November 9). Does the digital divide rhetoric do more harm than good? *The Chronicle of Higher Education, Information Technology.* http://chronicle.com/infotech.