Testing Common Sense

Robert J. Sternberg  
Yale University  
Richard K. Wagner  
Florida State University  
Wendy M. Williams  
Yale University  
Joseph A. Horvath  
Yale University

Predictors of success in school, such as conventional psychometric intelligence (e.g., IQ) tests, are less predictive of success out of school. Even the most charitable estimates of the relation between intelligence test scores and real-world criteria such as job performance indicate that approximately three fourths of the variance in real-world performance is not accounted for by intelligence test performance. Researchers have begun to explore new constructs in search of measures to supplement existing cognitive ability tests as predictors of real-world performance. Among the most promising constructs is practical intelligence, or common sense. Performance on measures of practical intelligence predicts real-world criteria such as job performance but is relatively unrelated to performance on intelligence tests and other common selection measures. Consequently, its contribution to prediction is largely independent of the contributions of existing measures, including measures of cognitive ability.

Each of us knows individuals who succeed in school but fail in their careers, or conversely, who fail in school but succeed in their careers. We have watched as graduate students, at the top of their class in the initial years of structured coursework, fall by the wayside when they must work independently on research and a dissertation. Most of us know of colleagues whose brilliance in their academic fields is matched only by their incompetence in social interactions. There are any number of reasons why someone might succeed in one environment but fail in another. A growing body of research, summarized in this article, suggests that differential success in academic and nonacademic environments reflects, in part, differences in the intellectual competencies required for success in these arenas.

Academic psychologists enjoy few perks, but one is the occasional opportunity to be thankful about their career choice. Consider the plight of the garbage collector, particularly in Tallahassee, Florida. If it is not enough that the summer heat makes any outdoor work unbearable, the city of Tallahassee adds insult to injury. Priding itself on the service it provides its citizens, Tallahassee requires physical labor far beyond the ordinary lifting and tossing of the standard-sized garbage cans placed carefully at curbside in other cities. In Tallahassee, each household fills a huge, city-issued trash container kept in its backyard. Trash collectors are required to locate and retrieve each full container from each backyard, haul them into the truck, and then drag back the empty containers to each yard.

Many of the garbage collectors are young high school dropouts who, because of their lack of education, might not be expected to do well on intelligence tests. And on the surface, the job appears to be physically but not intellectually challenging. Each stop simply requires two trips to the backyard, one to retrieve the full can, another to replace it when empty. Or so we thought.

After observing this collection routine one summer, we noticed that a new, older man joined the crew, and the routine changed. The change involved relaxing the constraint that each household retain the same garbage container. After all, the trash bins are identical, issued by the city rather than purchased with personal funds. The new routine consisted of wheeling the last house’s empty can into the current house’s backyard, leaving it to replace the full can that was in turn wheeled to the truck to be emptied. Once emptied, this can was now wheeled to the backyard of the next household to replace its full can, and so on. What had required two trips back and forth to each house now required only one. The new man’s insight cut the work nearly in half.

What kind of intelligence enables a person to come up with this kind of strategy for reducing effort by half, a strategy that had eluded well-educated observers such as the present authors, other garbage collectors, and the managers who trained them? And how well is this kind of intelligence reflected in an IQ score? An anecdote by Seymour Sarason, a psychologist at Yale, provides little grounds for optimism. When he reported to his first job—administering intelligence tests at a school for the mentally retarded—he could not begin testing because the students had cleverly eluded elaborate security precautions and escaped. When the students were rounded up, Sarason proceeded to administer the Porteus Maze Test, a paper-and-pencil intelligence test that involves finding the way

Editor’s note. Joseph D. Matarazzo served as action editor for this article.

Author’s note. Robert J. Sternberg, Wendy M. Williams, and Joseph A. Horvath, Department of Psychology, Yale University; Richard K. Wagner, Department of Psychology, Florida State University, Tallahassee.

Correspondence concerning this article should be addressed to Robert J. Sternberg, Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205.

November 1995 • American Psychologist
Copyright 1995 by the American Psychological Association, Inc. 0003-066X/95/33-00
Vol. 30, No. 11, 912-917
corrections raises the average validity coefficient to the level of about .5 (Hunter & Hunter, 1984; Schmidt & Hunter, 1981). Of course, this validity coefficient represents a hypothetical level, not one that is routinely obtained in practice. But even if one adopts the more optimistic hypothetical figure of .5, intelligence test scores account for only 25% of the variance in job performance. Whether you view the glass as being one-quarter filled or three-quarters empty, room for improvement exists. Researchers have therefore begun to explore new constructs in search of measures to supplement existing cognitive ability tests. Among the most promising constructs is practical intelligence, or common sense.

Practical Intelligence

Neisser (1976) was one of the first psychologists to press the distinction between academic and practical intelligence. Neisser described academic intelligence tasks (common in the classroom and on intelligence tests) as (a) formulated by others, (b) often of little or no intrinsic interest, (c) having all needed information available from the beginning, and (d) disembodied from an individual's ordinary experience. In addition, one should consider that these tasks (e) usually are well defined, (f) have but one correct answer, and (g) often have just one method of obtaining the correct solution (Wagner & Sternberg, 1985).

Note that these characteristics do not apply as well to many of the problems people face in their daily lives, including many of the problems at work. In direct contrast, work problems often are (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.

Laypersons have long recognized a distinction between academic intelligence (book smarts) and practical intelligence (street smarts). This distinction is represented in everyday parlance by expressions such as "learning the ropes" and "getting your feet wet." This distinction also figures prominently in the implicit theories of intelligence held by both laypeople and researchers. Sternberg, Conway, Kettou, and Benirschke (1981) asked samples of laypeople in a supermarket, a library, and a train station, as well as samples of academic researchers who study intelligence, to provide and rate the importance of characteristics of intelligent individuals. Factor analyses of the ratings supported a distinction between academic and practical aspects of intelligence for laypeople and experts alike.

Older adults commonly report growth in practical abilities over the years, even though their academic abilities decline. Williams, Denney, and Schadler (1983) interviewed men and women over the age of 65 about their perception of changes in their ability to think, reason, and solve problems as they aged. Although performance
on traditional cognitive ability measures typically peaks at the end of formal schooling, 76% of the older adults in the Williams et al. (1983) study believed that their ability to think, reason, and solve problems had actually increased over the years, with 20% reporting no change and only 4% reporting that their abilities had declined with age. When confronted with the fact of decline in psychometric test performance upon completion of formal schooling, the older sample said that they were talking about solving different kinds of problems than those found on cognitive ability tests—problems they referred to as "everyday" and "financial" problems.

Horn and Cattell (1966), in their theory of fluid and crystallized intelligence, provided a theoretical language with which to describe age-related changes in intellectual ability. According to their theory, fluid abilities are required to deal with novelty in the immediate testing situation (e.g., induction of the next letter in a letter series problem). Crystallized abilities reflect acculturated knowledge (e.g., the meaning of a low-frequency vocabulary word). A number of studies have shown that fluid abilities are vulnerable to age-related decline but that crystallized abilities are maintained throughout adulthood (Dixon & Baltes, 1985; Horn, 1982; Labouve-Vief, 1982; Schaie, 1977/1978).

Recall that practical problems are characterized by, among other things, an apparent absence of information necessary for a solution and for relevance to everyday experience. By contrast, academic problems are characterized by the presence, in the specification of a problem, of all the information necessary to solve the problem. Furthermore, academic problems are typically unrelated to an individual's ordinary experience. Thus, crystallized intelligence in the form of acculturated knowledge is more relevant to the solution of practical problems than it is to the solution of academic problems, at least as we are defining these terms. Conversely, fluid abilities, such as those required to solve letter series and figural analogy problems, are more relevant to the solution of academic problems than to the solution of practical problems. It follows that the growth in practical abilities that older participants report may reflect the greater contribution of maintained abilities—specifically, crystallized intelligence—in the solution of practical, everyday problems.

**Empirical Studies of Practical Intelligence**

The idea that practical and academic abilities follow different courses in adult development finds support in a variety of studies. For example, Denney and Palmer (1981) gave 84 adults between the ages of 20 and 79 years two types of reasoning problems: a traditional cognitive measure, the Twenty Questions Task (Mosher & Hornsby, 1966); and a problem-solving task involving real-life situations such as, "If you were traveling by car and got stranded out on an interstate highway during a blizzard, what would you do?" or, "Now let's assume that you lived in an apartment that didn't have any windows on the same side as the front door. Let's say that at 2:00 a.m. you heard a loud knock on the door and someone yelled, 'Open up. It's the police.' What would you do?"

The most interesting result of the Denney and Palmer study for the purposes of this article is 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 decreased linearly after age 20. Performance on the practical problem-solving task increased to a peak in the 40 and 50 year-old groups, then declined.

Cornelius and Caspi (1987) obtained similar results in a study of 126 adults between the ages of 20 and 78. They examined the relations between fluid intelligence, crystallized intelligence, and everyday problem solving. Cornelius and Caspi gave their participants traditional measures of fluid ability (Letter Series) and crystallized ability (Verbal Meanings), as well as an everyday problem-solving inventory that sampled the domains of consumer problems (a landlord who won't make repairs), information seeking (additional data is needed to fill out a complicated form), personal concerns (you want to attend a concert but are unsure whether it is safe to go), family problems (responding to criticism from a parent or child), problems with friends (getting a friend to visit you more often), and work problems (you were passed over for a promotion).

The measure of crystallized ability was given to determine whether the development of everyday problem solving was more similar to the development of crystallized ability than to the development of fluid ability. Performance on the measure of fluid ability 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 ability increased through age 70. Although Cornelius and Caspi's
are as


tion

lifes

The

lar

ive

by,

got

on

md.

edy

ch

its

ge
ed

er

lrea

me

's


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 intelligence test scores, arithmetic test scores, and grades.

Ceci and Liker (1986, 1988) carried out a study of expert racetrack handicappers. They studied strategies used by handicappers to predict post time odds at the racetrack. Expert handicappers used a highly complex algorithm for predicting post time odds that involved interactions among seven kinds of information. One obvious piece of useful information was a horse's speed on a previous outing. 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 other horses passed and where the attempted passes took place. These adjustments are important because they affect how much of the race is run away from the rail. By adjusting posted times for these factors, a better measure of a horse's speed is obtained. Use of the complex interaction in prediction would seem to require considerable cognitive ability (at least as it is traditionally measured). However, Ceci and Liker reported that the degree to which a handicapper used the interaction (determined by the regression weight for this term in a multiple regression of the handicappers' predicted odds) was unrelated to the handicapper's IQ, \( M = 97, r = -.07, p > .05 \).

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, & 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 $0.98 or 24 ounces for $2.29. One might adopt the strategy of always buying the largest

(1987) participants showed peak performance later in life than did Denney and Palmer's (1981), the pattern of traditional cognitive task performance peaking sooner than practical task performance was consistent across the studies. In addition to the developmental function of task performance, Cornelius and Caspi also examined the relation between performance on the fluid-ability and everyday problem-solving tasks and reported a modest correlation between the tasks \( r = .29, p < .01 \). The correlation between everyday problem-solving ability and crystallized ability was not higher \( r = .27, p < .01 \), leading Cornelius and Caspi to conclude that everyday problem solving was not reducible to crystallized ability, despite their similar developmental functions.

In summary, there is reason to believe that, whereas the ability to solve strictly academic problems declines from early to late adulthood, the ability to solve problems of a practical nature is maintained or even increased through late adulthood. The available evidence suggests that older individuals compensate for declining fluid abilities by restricting their domains of activity to those they know well (Baltes & Baltes, 1990) and by applying specialized procedural and declarative knowledge. For example, Salthouse (1984) has shown that age-related decrements at the "molecular" level (e.g., speed in 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).

These findings imply that, because fluid abilities are weaker determinants of performance on practical problems than they are of performance on academic problems, the use of scores on traditional cognitive ability tests to predict real-world performance should be problematic. In the past decade, a number of studies have addressed this and related issues. These studies, carried out in a wide range of settings and cultures, have been summarized
size, assuming that the largest size is always the most economical. However, the researchers (and savvy shoppers) learned that the largest 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 (though not completely accurate) enough to determine which size to buy. As for the oatmeal example, the kind of strategy used by effective shoppers was to recognize that 10 ounces for $0.98 is about 10 cents per ounce, and at that price, 24 ounces would be about $2.40, as opposed to the actual price of $2.29. Another common strategy involved mentally changing a size and price 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 $0.98 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, but that are much more easily computed in the absence of a calculator.

Another result of interest was that when the shoppers were given the M.I.T. mental arithmetic test, no relation was found between test performance and accuracy in picking the best values (Lave, Murtaugh, & de la Roche, 1984; Murtaugh, 1985). The same principle that applies to adults appears also to apply to children: Carraher, Carraher, and Schliemann (1985) found that Brazilian street children who could apply sophisticated mathematical strategies in their street vending were unable to do the same in a classroom setting.

One more example of a study of everyday mathematics was provided by individuals asked to play the role of city managers for the computer-simulated city of Lohhausien (Dörner & Kreuzig, 1983; Dörner, Kreuzig, Reither, & Staudel, 1983). A variety of problems were presented to these individuals, such as how to best 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 IQ 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 IQ and complexity of strategies employed.

**Tacit Knowledge**

The distinction between academic and practical kinds of intelligence is paralleled by a similar distinction between two types of knowledge (Sternberg & Caruso, 1985; Wagner, 1987; Wagner & Sternberg, 1985, 1986). An academically intelligent individual is characterized by their facile acquisition of *formal academic knowledge*, the knowledge sampled by the ubiquitous intelligence tests and related aptitude tests. Conversely, the hallmark of the practically intelligent individual is their facile acquisition and use of *tacit knowledge*. Tacit knowledge refers to action-oriented knowledge, acquired without direct help from others, that allows individuals to achieve goals they personally value (Horvath et al., in press). The acquisition and use of such knowledge appears to be uniquely important to competent performance in real-world endeavors. In this section we discuss the characteristic features of tacit knowledge, describe methods used in the testing of tacit knowledge, and review empirical support for the tacit knowledge construct.

**What Is Tacit Knowledge?**

There are three characteristic features of tacit knowledge. These features address, respectively, the structure of tacit knowledge, the conditions of its use, and the conditions under which it is acquired. First, tacit knowledge is procedural in nature. Second, tacit knowledge is relevant to the attainment of goals people value. Third, tacit knowledge is acquired with little help from others. Knowledge containing these three properties is called tacit because it often must be inferred from actions or statements. Please note, however, that although we have used the term *tacit* to refer to this type of knowledge, the intention or content of the tacit knowledge concept is not fully captured by the meaning of the lexical item *tacit*. Tacit knowledge is typically implied rather than stated explicitly—but there is more to the tacit knowledge concept than this most salient feature.

T tacit knowledge is procedural. Tacit knowledge is intimately related to action. It takes the form of "knowing how" rather than "knowing that" (Ryle, 1949). This sort
of knowledge (knowing how) is called *procedural knowledge*, and it is contrasted with *declarative knowledge* (knowing that). More precisely, procedural knowledge is knowledge represented in a way that commits it to a particular use or set of uses (Winograd, 1975). Procedural knowledge can be represented, formally, as condition-action pairs of the general form:

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 can be quite complex. In fact, much of the tacit knowledge that we have observed seems to take the form of complex, multiconditional rules for how to pursue particular goals in particular situations. For example, knowledge about getting along with one’s superior might be represented in a form with a compound condition:

IF (you need to deliver bad news)

AND

IF (it is Monday morning)

AND

IF (the boss’s golf game was rained out the day before)

AND

IF (the staff seems to be “walking on eggs”)

THEN (wait until later)

As this example suggests, tacit knowledge is always wedded to particular uses in particular situations or in classes of situations. Individuals who are queried about their knowledge will often begin by articulating general rules in roughly declarative form (e.g., “a good leader needs to know people”). When such general statements are probed, however, they often reveal themselves to be abstract or summary representations for a family of complex specified procedural rules (e.g., rules about how to judge people accurately for a variety of purposes and under a variety of circumstances). Thus, procedural structure is characteristic of tacit knowledge.

Tacit knowledge is practically useful. Tacit knowledge is instrumental to the attainment of goals people value. The more highly valued a goal is, and the more directly the knowledge supports the attainment of the goal, the more useful is the knowledge. For example, knowledge about how to make subordinates feel valued is practically useful for managers or leaders who value that outcome, but is not practically useful for those who are unconcerned with making their subordinates feel valued. Thus, tacit knowledge is distinguished from knowledge, even “how to” knowledge, that is irrelevant to goals that people care about personally.

Tacit knowledge is acquired without direct help from others. Tacit knowledge is usually acquired on one’s own. It is tacit knowledge that is untaught, underemphasized, or poorly conveyed relative to its importance for practical success. Thus, tacit knowledge is acquired under conditions of minimal environmental support. Environmental support refers to either people or media that help the individual acquire knowledge. When people or media support the acquisition of knowledge, they facilitate three knowledge acquisition components: selective encoding, selective combination, and selective comparison (Sternberg, 1983a, 1988). That is, when an individual is helped to distinguish more from less important information, is helped to combine elements of knowledge in useful ways, and is helped to identify knowledge in memory that may be useful in the present, then the individual has been supported in acquiring new knowledge. To the extent that this help is absent, the individual has not been supported.

To review, there are three characteristic features of tacit knowledge: (a) procedural structure, (b) high usefulness, and (c) low environmental support for acquisition. An important part of what makes the tacit knowledge concept a coherent one is the fact that these features are related to one another in nonarbitrary ways. In other words, we can explain why these features go together in the specification of a natural category of knowledge. We believe that this explanation strengthens the argument that tacit knowledge should be considered a well-formed concept.

First, it makes sense that procedural structure and high usefulness should both characterize a natural category of knowledge. Proceduralized knowledge also tends to be practically useful because it contains within it the specification of how it is used. Declarative knowledge, by contrast, is nonspecific with respect to use and, as a consequence, may remain unused or inert. Thus, procedural knowledge is more likely (than knowledge otherwise structured) to be instrumentally relevant in the pursuit of personally valued goals.

It also makes sense that high usefulness and low environmental support should both characterize a natural category of knowledge. Knowledge acquired in the face of low environmental support often confers a comparative advantage and thus tends to be practically useful in a competitive environment. When knowledge must be acquired in the face of low environmental support, the probability that some individuals will fail to acquire it increases. When some individuals fail to acquire knowledge, others who succeed in acquiring the knowledge may gain a competitive advantage over those who fail to acquire it. Note that the magnitude of this advantage would be lower if the knowledge in question was highly supported by the environment (i.e., explicitly and effectively taught), because more people would be expected to acquire and use it. Because many of the goals that individuals personally value are pursued in competition with...
other people, one may speculate that knowledge acquired under conditions of low environmental support is often particularly useful. This knowledge is more likely to differentiate individuals than is highly supported knowledge.

Finally, it makes sense that low environmental support and procedural structure should both characterize a natural category of knowledge. Proceduralized knowledge is often difficult to articulate and, thus, is more likely to be omitted from discussion or poorly conveyed. People know more than they can easily tell, and procedural knowledge is often especially difficult to articulate. Furthermore, procedural knowledge may become so highly automatized that people lose access to it completely. For these reasons, procedural knowledge is more likely than declarative knowledge to be acquired under conditions of low environmental support.

This discussion suggests that there is more to the tacit knowledge concept than a set of features assembled ad hoc to explain regularities in correlational data. Rather, the tacit knowledge concept is a coherent one, described not simply by a set of characteristic features but also by a set of nonarbitrary relations among those features.

**Testing Tacit Knowledge**

**Instruments.** Researchers have shown that tacit knowledge can be effectively measured (Sternberg, Wagner, & Okagaki, 1993; Wagner, 1987; Wagner & Sternberg, 1985, 1991; Williams & Sternberg, in press). The measurement instruments typically used consist of a set of work-related situations, each with between 5 and 20 response items. Each situation poses a problem for the participant to solve, and the participant indicates how he or she would solve the problem by rating the various response items. For example, in a hypothetical situation presented to a business manager, a subordinate whom the manager does not know well has come to him for advice on how to succeed in business. The manager is asked to rate each of several factors (usually on a 1 = low to 9 = high scale), according to their importance for succeeding in the company. Examples of factors might include (a) setting priorities that reflect the importance of each task, (b) trying always to work on what one is in the mood to do, and (c) doing routine tasks early in the day to make sure they are completed. Additional examples of work-related situations and associated response items are given in the Appendix.

Similarly, the tacit knowledge measurement instrument developed by Williams and Sternberg (in press) contains statements describing actions taken in the workplace, which participants rate for how characteristic the actions are of their behavior. In addition, complex open-ended problem situations are described, and participants are asked to write plans of action that show how they would handle the situations.

**Scoring.** The procedure for scoring a tacit knowledge test has evolved across several studies, and various scoring approaches are briefly described here. In Wagner and Sternberg’s (1985b) study, the tacit knowledge test was scored by correlating ratings on each response item with a dummy variable representing group membership (e.g., 3 = experienced manager, 2 = business school student, 1 = undergraduate). A positive correlation between item and group membership indicated that higher ratings were associated with greater levels of expertise in the domain, whereas a negative correlation indicated that higher ratings were associated with lower levels of expertise in the domain. Items showing significant item-group correlations were retained for further analysis. Ratings for these items were summed across items in a given subscale, and these summed values served as predictor variables in analyzing the relationship, within groups, between tacit knowledge and job performance.

A second procedure for scoring tacit knowledge tests was employed by Wagner (1987). A sample of practically intelligent individuals (this time, academic psychologists) was obtained through a nomination process. The tacit knowledge test was administered to these individuals, and an expert profile was generated that represented the central tendency of their responses. Tacit knowledge tests for participants were scored, separately for each item subscale, as the sum of their squared deviations from this expert profile. Note that this scoring method, unlike that described previously, allows for meaningful comparisons between groups.

A third procedure for scoring tacit knowledge tests was that of Wagner, Rashotte, and Sternberg (1992). In a study of tacit knowledge for sales, they collected rules of thumb through reading and interviews. According to its dictionary definition, a rule of thumb is "a useful principle with wide application, not intended to be strictly accurate" (Morris, 1987, p. 1134). Examples of rules of thumb that differentiated expert from novice salespersons included, "Penetrate smoke screens by asking ‘what if . . . ?’ questions," and "In evaluating your success, think in terms of tasks accomplished rather than hours spent working." These rules of thumb were grouped into categories and used to generate a set of work-related situations. Response items were constructed so that some items represented correct application of the rules of thumb, whereas other items represented incorrect or distorted application of the rules of thumb. The tacit knowledge test was scored for the degree to which participants preferred response items that represented correct applications of the rules of thumb.

**Findings From the Tacit Knowledge Research Program**

Earlier we described several studies in which participants of different ages were given measures of everyday problem solving and measures of traditional cognitive abilities (Cornelius & Caspi, 1987; Denney & Palmer, 1981). The results suggested different developmental functions for the two kinds of abilities: Performance on traditional cognitive ability measures peaked in early adulthood, but performance on everyday problem-solving measures continued to improve through later adulthood. Which of these age-related changes in intelligence is best described as true cognitive development?
these two functions better characterizes the development of tacit knowledge?

In a cross-sectional study, we administered a tacit knowledge inventory to three groups of participants, totaling 127 individuals, who differed in their breadth of experience and formal training in business management (Wagner & Sternberg, 1985, Experiment 2). One group consisted of 54 business managers, another group consisted of 51 business school graduate students, and a third group consisted of 22 Yale undergraduates. The means and standard deviations for amount of managerial experience were 16.6 (9.9) years for the business manager group; 2.2 (2.5) for the business graduate student group; and 0.0 (0.0) for the undergraduate group. Group differences were found on 39 of the response item ratings, with a binomial test of the probability of finding this many significant differences by chance yielding p < .0001. We conclude from this study that there were genuine differences in the ratings for the groups. We obtained comparable results for academic psychologists (Wagner & Sternberg, 1985a, Experiment 1). In a second cross-sectional study, we obtained tacit knowledge scores from three new groups of 64 managers, 25 business graduate students, and 60 Yale undergraduates (Wagner, 1987), and we used a prototype-based scoring system that allowed direct comparisons of the performance of the three groups. In this study, the business managers group, whose average age was 50, outperformed the business graduate students and the undergraduates. The business graduate students in turn outperformed the undergraduates. Again, comparable results were obtained for psychology professors, psychology graduate students, and undergraduates. Although these studies did not sample different age ranges as exhaustively as in the studies previously described (Cornelius & Caspi, 1987; Denney & Palmer 1981), the results suggested that the development of tacit knowledge more closely resembles the development of everyday problem solving than that of cognitive ability as traditionally measured.

In a later study that focused on the development of tacit knowledge over the managerial career, Williams and Sternberg (in press) used extensive interviews and observations to construct both a general and a level-specific tacit knowledge measure. We administered this measure to all executives in four high technology manufacturing companies. We also obtained nominations from managers' superiors for "outstanding" and "underperforming" managers at the lower, middle, and upper levels. This approach enabled us 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.

Our results showed that there was indeed specialized tacit knowledge for each of the three management levels and that this knowledge was differentially related to success. We derived these results by 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 was substantially more important to upper-level executives than to middle- or lower-level executives. Knowledge about maintaining appropriate levels of control became progressively more significant at higher levels of management. Knowledge about self-motivation, self-direction, self-awareness, and personal organization was roughly comparable in importance at the lower and middle levels, and became somewhat more important at the upper level. Finally, knowledge about completing tasks and working effectively within the business environment was 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 was essential for all managers, but especially for those in the upper level. Knowledge about supporting, cooperating with, and understanding others was extremely important for upper-level executives, very important for middle-level executives, and somewhat important for lower-level executives.

**Questions About the Tacit Knowledge Construct**

We have argued elsewhere that the "g-oecntric view" of intelligence and job performance is wrong—that there is more to successfully predicting job performance than just measuring the general factor from conventional psychometric tests of intelligence (see Sternberg & Wagner, 1993). We suggested an aspect of practical intelligence, tacit knowledge, as a key ingredient to job success. Not everyone has agreed with this point of view. Jensen (1993), Schmidt and Hunter (1993), and Rec and Earles (1993) have presented various arguments against this position. This section addresses questions raised by critics of the tacit knowledge research program.

Are individual differences in tacit knowledge domain-general? Are measures of tacit knowledge highly domain-specific tests of job knowledge, analogous to a test for mechanics that requires identifying a crescent wrench, or do they represent some more general construct? The evidence to date is more compatible with the view that measures of tacit knowledge assess a relatively general construct.

Two kinds of factor analysis were performed on the tacit knowledge scores of a sample of 64 business managers (Wagner, 1987). A principal components analysis yielded a first principal component that accounted for 44% of the total variance, and for 76% of total variance after the correlations among scores were disattenuated for unreliability. The residual matrix was not significant after extracting the first principal component. A first principal component accounting for about 40% of total variance is typical of analyses carried out on traditional cognitive ability subtests. A confirmatory factor analysis was performed to test alternative models of the factor
structure of the tacit knowledge inventory more formally. The results supported the generality of tacit knowledge. A model consisting of a single general factor provided the best fit to the data and yielded small and nonsignificant differences between predicted and observed covariances. The root mean square residual was .08 (N = 64, $\chi^2(9) = 12.13, p > .05$).

The domain generality of tacit knowledge was given additional support when the identical tacit knowledge framework was used to construct a new measure of tacit knowledge for the domain of academic psychology. A parallel study that included samples of psychology professors, graduate students, and undergraduates yielded a pattern of results nearly identical to that found in business samples. More importantly, a group of 60 undergraduates was given tacit knowledge measures for both domains—business management and academic psychology—in counterbalanced order. After determining that order of administration did not affect the latent structure of the two tacit knowledge measures, we calculated correlations between scores across measures. The magnitude of these cross-domain correlations was .58 for total score, .52 for managing oneself, .47 for managing tasks, and .52 for managing others (components of our tacit knowledge construct), all significant at the $p < .001$ level. These results support the domain generality of individual differences in tacit knowledge.

Are tacit knowledge inventories just intelligence tests in disguise? If individual differences in tacit knowledge appear to have some domain generality, has this accidentally reinvented the concept of "g," or general ability, that can be extracted from an intelligence test? Results from several studies of tacit knowledge, in which participants have been given a traditional measure of cognitive ability in addition to a tacit knowledge inventory, suggest that this is not the case.

For example, Wagner and Sternberg (1985) gave the Verbal Reasoning subtest of the Differential Aptitude Tests (Form T) to a sample of 22 undergraduates. The correlation between tacit knowledge and verbal reasoning was .16 ($p > .05$). In subsequent studies, a deviation scoring system, in which lower scores indicated better performance than higher scores, was used to quantify tacit knowledge. Thus, a positive relation between tacit knowledge and cognitive ability would be represented by a negative correlation. For a sample of 60 undergraduates, the correlation between tacit knowledge and verbal reasoning was $-.12 (p > .05)$.

One important limitation of these results is that the participants were Yale undergraduates and thus represented a restricted range of verbal ability. In addition, undergraduates have relatively little tacit knowledge compared with experienced managers. Rather different correlations between tacit knowledge and IQ might therefore be expected for other groups, such as business managers. We administered the Tacit Knowledge Inventory for Managers to a sample of 45 managers who participated in a leadership development program at the Center for Creative Leadership in Greensboro, North Carolina (Wagner & Sternberg, 1990). Participants routinely completed a battery of tests, including an intelligence test. For this sample, the correlation between tacit knowledge and IQ was $-.14 (p > .05)$.

But even business managers represent a restricted range in IQ, and perhaps in tacit knowledge as well. What would be the relation between tacit knowledge and IQ in a more general sample? In a study at the Human Resources Laboratory at Brooks Air Force Base that was supervised by Malcolm Ree, Eddy (1988) examined relations between the Tacit Knowledge Inventory for Managers and the Armed Services Vocational Aptitude Battery (ASVAB) for a sample of 631 Air Force Recruits, 29% of whom were women, and 19% of whom were members of a minority group. The ASVAB is a multiple-aptitude battery used for the selection of candidates by all branches of the United States Armed Forces. Prior studies of the ASVAB suggested that it is a typical measure of cognitive ability, with correlations between ASVAB scores and other cognitive ability measures of about .7. Factor analytic studies have also suggested that the ASVAB appears to measure the same verbal, quantitative, and mechanical abilities as the Differential Aptitude Tests, and measures 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 2 correlations were significantly different from 0, 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 ability, and mathematics) and a distinct tacit knowledge factor. The factor loading for the Tacit Knowledge Inventory for Managers score on the tacit knowledge factor was $-.99$, with a maximum loading for the score 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 0 ($-.075, .003, .096$, and $.082$).

One final 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, although their performance on traditional selection measures, such as intelligence tests, does not. Eddy (1988) did not report scores separately by race and sex, but did report correlations between scores and dummy variables that indicated race and sex. Significant correlations in the $-.2$ to $-.4$ range between ASVAB subtest scores and both race and sex indicated that on the ASVAB, minority group members had poorer scores than majority group members, and women scored lower than men. However, nonsignificant correlations between tacit knowledge and both race (.03) and sex (.02) indicated
comparable levels of performance on the tacit knowledge measures between minority and majority group members and between women and men.

Does performance on measures of tacit knowledge uniquely predict performance in management? In several early studies, we gave our tacit knowledge measure to samples of business managers and examined correlations between tacit knowledge scores and criterion reference measures of performance in business. For example, in a sample of 54 (Wagner & Sternberg, 1985) and 64 (Wagner, 1987) business managers, we found correlations ranging from .2 to .4 between tacit knowledge score and criteria such as salary, years of management experience, and whether the manager worked for a company at the top of the Fortune 500 list. These uncorrected correlations were in the range of the average correlation between cognitive ability test scores and job performance of .2 (Wigdor & Garner, 1982).

In these studies, the managers were from a wide range of companies, and only global criterion measures—such as salary and years of management experience—were available for study. When more precise criterion measures have been available, higher correlations between tacit knowledge and performance have been found. For example, in a study of bank branch managers (Wagner & Sternberg, 1985), the correlation between tacit knowledge and average percentage of merit-based salary increase was .48 (p < .05). The correlation between tacit knowledge and average performance rating for the category of "generating new business for the bank" was .56 (p < .05).

Further support for the predictive validity of tacit knowledge measures was provided by the previously mentioned study of business managers who participated in the Leadership Development Program at the Center for Creative Leadership (Wagner & Sternberg, 1990). In this study, we were able to examine correlations among a variety of measures, including the Tacit Knowledge Inventory for Managers. The appropriate statistic to determine what will be gained by adding a test to existing selection procedures, or conversely, what will be lost by deleting a test, is the squared semipartial correlation coefficient, or change in $R^2$ from hierarchical regression analyses. We were able to provide an empirical demonstration of this type of validity assessment in the Center for Creative Leadership study.

Every manager who participated in the Leadership Development Program at the Center for Creative Leadership completed a battery of tests. By adding the Tacit Knowledge Inventory for Managers to the battery, we were able to determine the unique predictive power of the inventory in the context of other measures commonly used in managerial selection. These measures included the Shipley Institute for Living Scale, an intelligence test; 17 subtest scores from the California Psychological Inventory, a self-report personality inventory; 6 subtest scores from 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; 4 subtest scores from the Myers—Briggs Type Indicator, a measure of cognitive style; the Kirton Adaptation Innovation Inventory, a measure of preference for innovation; and 5 subtest scores from the Managerial Job Satisfaction Questionnaire, a measure of job satisfaction.

The criterion measure of managerial performance was behavioral assessment data ratings in two small-group managerial simulations called Earth II and Energy International. The managers worked in groups of five to solve realistic business problems. Trained observers rated the performance of the managers in eight categories: activity level, discussion leading, influencing others, problem analysis, task orientation, motivating others, verbal effectiveness, and interpersonal skills. To obtain a criterion measure with sufficient reliability, the ratings were averaged and summed across the two simulations. The Spearman—Brown corrected split-half reliability of this total score was .59.

Beginning with zero-order correlations, the best predictors of the criterion score of managerial performance were tacit knowledge (r = —.61, p < .001) and IQ (r = .38, p < .001). (The negative correlation for tacit knowledge was expected because of 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 IQ was not significantly different from 0 (r = —.14, p > .05). We carried out a series of hierarchical regressions to examine the unique predictive value of tacit knowledge when used in conjunction with existing measures. For each hierarchical regression analysis, the unique prediction of the Tacit Knowledge Inventory for Managers was represented by the change in $R^2$ from a restricted model to a full model. In each case, the restricted model contained various measures, and the full model was created by adding the Tacit Knowledge Inventory for Managers as another predictor. If adding the tacit knowledge score resulted in a significant and substantial change in $R^2$, one could conclude that the predictive relation between tacit knowledge and the criterion measure was not subsumed by the set of predictors in the restricted model. The results are presented in Table 1.

In Table 1, the measures listed in the column titled "Measures in Restricted Model" were the predictors that had already been entered in the regression before entering the tacit knowledge score. In the first example, the sole predictor used in the restricted model was IQ. The values reported in the column titled "$R^2$ Change When Tacit Knowledge Is Added" are the increases in variance accounted for in criterion when tacit knowledge was added to the prediction equation. For the first example, tacit knowledge accounts for an additional 32% of criterion variance that is not accounted for by IQ. The values reported in the column titled "$R^2$ for Full Model" indicate the proportions of variance in the criterion that is accounted for by tacit knowledge and the other measures when used in conjunction.
Table 1
Hierarchical Regression Results From the Center for Creative Leadership Study

<table>
<thead>
<tr>
<th>Measures in restricted model</th>
<th>$R^2$ change when tacit knowledge is added</th>
<th>$R$ for full model</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>-.32***</td>
<td>.40***</td>
</tr>
<tr>
<td>Seventeen CPI subtests, IQ</td>
<td>-.22**</td>
<td>.66*</td>
</tr>
<tr>
<td>6 FIO-B subtests, IQ</td>
<td>-.32***</td>
<td>.65***</td>
</tr>
<tr>
<td>Field Independence, IQ</td>
<td>-.28***</td>
<td>.47***</td>
</tr>
<tr>
<td>Kirton innovation, IQ</td>
<td>-.33***</td>
<td>.50***</td>
</tr>
<tr>
<td>Four Myers-Briggs subtests, IQ</td>
<td>-.35***</td>
<td>.56***</td>
</tr>
<tr>
<td>Five Job Satisfaction subtests, IQ</td>
<td>-.32**</td>
<td>.57***</td>
</tr>
</tbody>
</table>

Note. CPI = California Psychological Inventory; FIO-B = Fundamental Interpersonal Relations Orientation-B.
*p < .05; **p < .01; ***p < .001.

In every case, tacit knowledge accounted for substantial and significant increases in variance. In addition, when tacit knowledge, IQ, and selected subtests from the personality inventories were combined as predictors, we accounted for nearly all of the reliable variance in the criterion. 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 be readily subsumed by the existing constructs of cognitive ability and personality represented by the other measures used in the study.

Williams and Sternberg (in press) also studied the interrelationship of tacit knowledge for management with demographic and experiential variables. (In this research, tacit knowledge was defined as the sum of squared deviations of participants’ ratings from nominated experts’ score arrays on a tacit knowledge measure.) We 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 ($r = .36, p < .001$). Note that these correlations were computed after controlling for background and educational experience. Tacit knowledge was also weakly associated with enhanced job satisfaction ($r = .23, p < .05$). Demographic and education variables unrelated to tacit knowledge included age, years of management experience, years in current position, degrees received, mother’s and father’s occupations, mother’s and father’s educational level attained, and mother’s and father’s degrees received. (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 suggesting the possibility that “deadwood” managers often stayed around a long time. 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 IQ, suggest that tacit knowledge overlaps with the portion of these measures that is not predicted by IQ.

This pattern of interrelationships between tacit knowledge scores and demographic and background variables prompted us to examine the prediction of our success measures, through hierarchical regression. These analyses showed whether tacit knowledge contained independent information related to success—information 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 to enter the regression equation was years of education, accounting for 19% of the variance ($p < .001$). The second variable was years of management experience, accounting for an additional 13% of the variance ($p < .001$). The third and final variable to enter was tacit knowledge, accounting for an additional 4% of the variance ($p = .04$), raising the total explained variance to 36%. In the regression predicting maximum compensation controlled for age, years of education entered the equation first, accounting for 27% of the variance ($p < .001$). Finally, tacit knowledge entered, explaining an additional 5% of the variance ($p = .03$). This regression may be viewed as the most significant, insofar as it demonstrated the value of tacit knowledge to managers who were relatively successful for their age.

The general conclusions to be drawn from all of the regression analyses are, first, that it is difficult to predict success measures, such as salary and maximum compensation, presumably because of the myriad effects upon such variables that were outside of the focus of this study. Nonetheless, approximately 40% of the variance in the success measures used in this study was explicable. For all four success measures, the educational variable was the most important, followed—in the case of salary and maximum compensation—by an experiential variable (years of management experience). After education and experience were included in the equations, tacit knowledge still explained a significant proportion of the variance in success. Thus, tacit knowledge contains information relevant to the prediction of success that is independent of that represented by the background and demographic variables.

Is tacit knowledge only important in business? Although our focus has been on the tacit knowledge of business managers, there is evidence that the construct also explains performance in other domains. In two studies of the tacit knowledge of academic psychology professors, correlations in the .4 to .5 range were found between tacit knowledge and criterion measures such as number of citations reported in the Social Science Citation Index and the Facets of the American Psychological Association.
the rated scholarly quality of an individual's departmental faculty (Wagner, 1987; Wagner & Sternberg, 1985). More recently, we have begun to investigate the role of tacit knowledge in the domain of sales (Wagner, Rashotte, & Sternberg, 1992). We have found correlations in the .3 to .4 range between measures of tacit knowledge about sales and criterion measures such as sales volume and sales awards received for a sample of life insurance salespersons. In this work, we also have been able to express the tacit knowledge of salespersons in terms of rules of thumb that serve as rough guides to action in sales situations. Expressing tacit knowledge in terms of rules of thumb may permit explicit training of at least some aspect of tacit knowledge. A preliminary training study in which undergraduates were trained in tacit knowledge relevant to the profession of sales found greater pretest-posttest differences in tacit knowledge for groups whose training identified relevant rules of thumb than for those whose training did not make any such identifications (Sternberg, Wagner, & Okagaki, 1993).

We have also studied the role of tacit knowledge in school performance (Sternberg, Okagaki, & Jackson, 1990; Williams et al., in press). A six-year program of research, called the Practical Intelligence for Schools Project, involved intensive observations and interviews of students and teachers to determine the tacit knowledge necessary for success in school. Curricula designed to train the essential tacit knowledge were developed and evaluated in matched-group controlled studies in schools across Connecticut. This work was undertaken in collaboration with Howard Gardner and with other researchers at Harvard University, who also developed and evaluated curricular materials in Massachusetts schools. A final composite Practical Intelligence for School (PIFS) curriculum has been created by the Yale and Harvard teams (Williams et al., in press) and is now being used in hundreds of classrooms across the United States and abroad.

The results of PIFS curriculum evaluations have been uniformly positive. In 1992–1993, Connecticut-area students receiving PIFS showed significantly greater increases in reading, writing, homework, and test-taking ability over the school year, compared with students in the same schools not receiving the curriculum (ANCOVA F for PIFS variable = 60.89, p < .0001). Furthermore, teachers, students, and administrators reported fewer behavioral problems in PIFS classes. This research demonstrated that tacit knowledge is instrumental to school success and, significantly, that it can be effectively and efficiently taught.

Conclusions

Approximately 20 years ago, McClelland (1973) questioned the validity of cognitive ability testing for predicting real-world criteria such as job performance, arguing in favor of competency tests that more closely reflect job performance itself. Subsequent reviews of the literature on the predictive validity of intelligence tests suggest that McClelland may have been pessimistic about the validity of intelligence tests. Individual differences in intelligence-test performance account for between 4% and 25% of the variance in such real-world criteria as job performance (Barrett & Depinet, 1991; Hunter & Hunter, 1984; Schmidt & Hunter, 1981; Wigdor & Garner, 1982). Nevertheless, between 75% and 96% of the variance in real-world criteria such as job performance cannot be accounted for by individual differences in intelligence test scores. We view the emerging literature on practical intelligence, or common sense, as a belated response to McClelland’s call for new methods to assess practical abilities. This literature provides three sources of evidence to support a distinction between academic and practical intelligence.

First, the distinction between academic and practical intelligence is entrenched in the conception of intelligence held by laypeople and researchers alike. In addition to evidence provided by studies of implicit theories of intelligence (Sternberg, 1985b; Sternberg, Conway, Kremen, & Benvenstein, 1981), analyses of researchers’ descriptions of the nature of intelligence suggest a prominent role for practical intelligence. 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 intelligence 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 intelligence theorists revealed concern about practical aspects of intelligence (Sternberg & 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 intelligence 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 intelligence is also supported by older adults’ perception of age-related changes in their ability to think and to solve problems (Williams, Denney, & Schadler, 1983). Three fourths of the older adults sampled believed that their ability 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 intelligence is the result of empirical studies of age-related changes in adults’ performance on academic and practical tasks. The results suggest different developmental functions for changes in performance on the two kinds of tasks across the adult life span. Whereas performance on intelligence tests, particularly those that measure fluid ability, begins to decline in middle adulthood, performance on measures of everyday problem solving continues to improve until old age (Cornelius & Caspi, 1987; Denney & Palmer, 1981; Horn & Cattell, 1966). Our own 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, Ruhotte, & Sternberg, 1992; Wagner & Sternberg, 1985; Williams & Sternberg, in press). These increases emerged despite probable decreases in intelligence test performance across groups, particularly for the manager studies.

The third source of evidence to support a distinction between academic and practical intelligence is the result of studies in which participants were assessed on both academic and practical tasks. The consistent result is little or no correlation between performance on the two kinds of tasks. IQ is 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 & Liker, 1986, 1988); (c) the complexity of strategies used in computer-simulated roles, such as city manager (Dörner & Kreuzig, 1983; Dörner, Kreuzig, Reither, & Staudel, 1983); and (d) the tacit knowledge of undergraduates (Wagner, 1987; Wagner & Sternberg, 1985), business managers (Wagner & Sternberg, 1990), salespeople (Wagner, Ruhotte, & Sternberg, 1992), and U.S. Air Force recruits (Eddy, 1988). In addition, the accuracy with which grocery shoppers identified quantities that provided the best value was unrelated to their performance on the M.I.T. mental arithmetic test (Lave, Murtaugh, & de la Roche, 1984; Murtaugh, 1985).

Our conclusions about the value of measures of academic and practical intelligence for predicting real-world performance differ from those of previous reviews of literature by McClelland (1973) and by Barrett and Depinet (1971). McClelland argued that measures of academic abilities are of little value in predicting real-world criteria such as job performance. Barrett and Depinet (1991) argued that there was little in measures of practical abilities for predicting job performance. Our view is that there is complementary value in both kinds of measures. We believe that differences between our conclusions and those of both McClelland and Barrett and Depinet derive from differences in the studies that were included in the reviews.

McClelland’s (1973) landmark article was published years before the emergence of meta-analysis, a methodology for cumulating results across studies. The results of meta-analytic studies, particularly when corrections for measurement error and restriction of range were used, provided larger estimates of the correlation between IQ and real-world criteria, such as job performance, than those apparent from inspection of the individual studies that were available to McClelland at the time of his review.

Although we agree with Barrett and Depinet’s (1991) conclusion that cognitive ability tests have some value for selection, we are at odds with their dismissal of measures of practical performance. In what they described as a comprehensive review of the relevant literature (p. 1012), Barrett and Depinet considered, among other issues, McClelland’s (1973) claim that practical tasks are necessary for predicting practical outcomes. In a section titled “Practical Tasks,” we were surprised that Barrett and Depinet reported the results of only a single recent study, by Willis and Schaeie (1986), concluding that it demonstrated that an “extremely high relationship existed between intelligence and performance on real-life tasks” (p. 1015). Barrett and Depinet ignored studies that did not support their thesis, even though the omitted studies (some of which we have discussed here) were described in chapters of the same book, Practical Intelligence (Sternberg & Wagner, 1986), from which they extracted Willis and Schaeie’s study. In fact, the study they included in their review was the sole study reported in Practical Intelligence that supported their thesis. Furthermore, from their description of the Willis and Schaeie study, one would not know that the criterion measure of performance on real-life tasks used in the study was, in fact, a paper-and-pencil psychometric test (the ETS Basic Skills Test; Educational Testing Service, 1977), with tasks such as writing paragraphs and describing the main theme, interpreting written guarantees for devices such as calculators, reading letters and determining on which points the authors are in agreement, and interpreting maps and charts. This test may measure basic skills relevant to real-world performance, but it is decidedly more academic than changing a flat tire or convincing your superiors to spend a million dollars on your idea.

Our concern about selective inclusion of studies is twofold. Obviously, selective inclusion of studies can result in biased conclusions. But it also discourages researchers from seeking generalizations that incorporate ostensibly disparate results. For example, by comparing the characteristics of the measures used by Willis and Schaeie (1986) with those used by other contributors, the following generalization emerged:

Looking across the studies reported in this volume, the correlation between measures of practical and academic intelligence varies as a function of the format of practical intelligence measure: Correlations are large when the practical intelligence measure is test-like, and virtually nonexistent when the practical intelligence measure is based on simulation. (Wagner, 1986, p. 372)

For the present and foreseeable future, we believe that the most viable approach to increasing the variance accounted for in real-world criteria (e.g., job performance) is to supplement existing intelligence and aptitude tests with selection of additional measures based on new constructs, such as practical intelligence. Although we are excited by the promise of a new generation of measures of practical intelligence, we are the first to admit that existing evidence for the new measures does not yet match that available for traditional cognitive–academic ability tests. However, a substantial amount of evidence indicates that performance on measures of practical intelligence is related to a wide variety of criterion measures of real-world performance, but relatively unrelated to traditional...
measures of academic intelligence. Consequently, the use of both kinds of measures results in more effective prediction than reliance on either kind alone.

REFERENCES


