

The Emperor's New Clothes: Additional Critiques of Sternberg's Practical Intelligence Theory

Michael A. McDaniel  
Virginia Commonwealth University

Deborah L. Whetzel  
Work Skills First, Inc.

Paper presented at the Fourth Annual Conference of the International Society for Intelligence Research, Newport Beach, CA. December, 2003. An earlier version of this paper was presented at the International Symposium on Emotional and Practical Intelligence. Berlin, Germany, July 2003.

Correspondence on this article may be addressed to Michael A. McDaniel  
E-mail: [mamcdani@vcu.edu](mailto:mamcdani@vcu.edu)

## Abstract

Gottfredson (2003) provided a detailed critique of Sternberg's (Sternberg, Forsythe, Hedlund, Horvath, Wagner, Williams, Snook & Grigorko, 2000) practical intelligence theory. Although very thorough, Gottfredson's work did not draw on relevant evidence from the situational judgment literature in industrial/organizational psychology. The current paper seeks to supplement Gottfredson's extensive critique of practical intelligence using research and practice in the situational judgment literature. This paper makes six major points. First, there is a research and practice tradition in personnel selection that has used the item type that Sternberg re-invented and called tacit knowledge. This item type is called situational judgment in the personnel selection field. There is an extensive literature on situational judgment tests that contradicts Sternberg's claims concerning practical intelligence. Second, we present evidence that such situational judgment tests do not measure a general factor, whether labeled practical intelligence, or something else. Both the individual test items and the tests are factorially complex and measure multiple known constructs. Third, these measures are moderately correlated with *g*. Fourth, these measures assess multiple constructs and are best viewed as measurement methods. In fact, the correlation of the measures with personality and cognitive constructs can be influenced by the instructions given to the respondents. Fifth, we compare the validity of situational judgment measures with *g* for predicting job performance. Sixth, we provide suggestions for future research on tests of this type.

## 1. Introduction

In their book, *Practical intelligence in everyday life*, Sternberg and his colleagues make two bold claims regarding the nature of practical or tacit intelligence. First, they claim that there exists a general factor of practical or tacit intelligence that is substantively distinct from *g*.

Second, they claim that this new construct, practical intelligence, predicts success in various domains as well as or better than *g*.

Sternberg's theory has a great deal of intuitive appeal because it offers an optimistic view, however, unsubstantiated by data. For example, the triarchic theory describes intelligence as something other than "book smarts" which suggests that even if one is not a good student or particularly bright academically, one can still be successful. Sternberg's triarchic theory of intelligence states that there are three broad abilities: analytical, creative, and practical. Thus, if one is low on one ability, one can still succeed on either of the other two. Intelligence theorists, on the other hand, have shown that individuals low on *g* are likely to be beset with a host of adverse societal outcomes, such as juvenile delinquency, adult crime, single parenthood, HIV infection, and poverty (Gordon, 1997). In the work domain, those low in *g* will tend to have poor job performance and are found in lower complexity, low paying jobs (Schmidt & Hunter, 1998). Sternberg's writings even offer hope to the aged with regard to the decline of fluid intelligence. He states that, "when the same individuals are followed across time in the framework of longitudinal design the findings show that, with respect to fluid intelligence, decline does not generally begin until the 60's . . ." (p. 39). This claim does not account for the vast amount of cross-sectional research showing that the decline in fluid intelligence starts in the mid-20's (Horn & Cattell, 1967).

Another reason for its appeal is the way it is described. Sternberg provided a great deal of anecdotal evidence to support his claims. For example, in an effort to distinguish between *g* and tacit intelligence, Sternberg cited the work of Ceci and Liker (1986) on racetrack handicappers and Lave (1988) on housewives shopping in supermarkets. These frequent references to "real people" make his claims seem plausible to the unsophisticated reader.

Gottfredson (2003) provided a thorough critique of Sternberg's claims by describing the volumes of research citing the importance of *g* and the paucity of research supporting Sternberg's theory. In detailing the extensive research base of *g*, she cites Carroll's "three stratum" theory in which each successive stratum in the hierarchy emerges from the common variance of the previous stratum. Stratum I includes narrow abilities, such as spatial relations, perceptual speed, and associative memory. Stratum II include broader factors, such as visual perception and general memory that combine many of the abilities in Stratum I. Stratum III consists of *g*, which is the only factor that is common to all Stratum II factors (Carroll, 1993). As described, *g* is the major component of all the moderately highly correlated Stratum II factors, which in turn are the major components of the Stratum I abilities. The "Gf-Gc theory" of fluid and crystallized intelligence also is encompassed by the three-stratum hierarchy (Cattell, 1987; Horn & Cattell, 1966). Fluid intelligence refers to the ability to solve cognitive problems immediately. Crystallized intelligence refers to general mental skills (i.e., language) that have been developed or "crystallized" from exercising fluid intelligence in the past. Fluid intelligence correlates nearly 1.0 with *g* and crystallized intelligence correlates approximately .8 with *g* (Gustafsson, 1988). As such, *g* is derived from or correlated with the most well known conceptualizations of intelligence to date.

In addition to citing the volumes of research on *g*, Gottfredson also noted the lack of empirical evidence supporting Sternberg's triarchic theory. Specifically, she states that the book, "... provides more of a collage of related theorizing than a carefully developed model of practical intelligence. And instead of collating into tables the data from two decade of research, the book gives the same unintegrated narrative summary of selected results study by study that has been published in similar form before" (Gottfredson, 2003, p. 345). Gottfredson provides a

detailed critique of six major empirical claims made by Sternberg and colleagues and finds little support for any of these claims. Given the length and detail provided by Gottfredson, repeating those arguments is beyond the scope of this paper. Readers are referred to the original work.

Brody (2003) presented an alternative theoretical analysis of research presented by Sternberg and his colleagues of studies designed to validate the available version of the Sternberg Triarchic Abilities Test (STAT) using students participating in a summer school program at Yale. The STAT measures three abilities, analytical, creative and practical using both multiple choice and essay questions. Using overall performance on tests, assignments, and final projects (academic achievement) as the dependent variables, their multiple regression analyses showed that practical ability had a significant independent contribution in only one of the 12 analyses. In fact, the prediction of academic achievement is only marginally improved using the combined influence of each of the triarchic abilities, as opposed to that derived from the analytical ability scores alone. These and other analyses provided by Brody suggest that the construct validity of the STAT for measuring the three triarchic abilities independent of *g* is questionable at best.

In sum, Brody's analyses questioned the measurement method used by Sternberg and colleagues and Gottfredson showed that there is one overall factor that defines intelligence—*g*.

## **2. Sources for this Critique**

The purpose of this paper is to supplement recent critiques of practical intelligence by drawing on research from industrial/organizational psychology, particularly personnel selection. This critique originates from several sources. First, there is the literature on the validity of *g* in personnel selection. Concerning the validity of *g* for predicting performance, Schmidt and Hunter (1998) summarized the meta-analytic evidence for the criterion-related validity (i.e., the

correlation between the predictor measure and job performance) of 19 types of predictors. They found that  $g$  was the second most valid predictor (.51) for jobs with moderate cognitive demands, which include approximately 60% of jobs in the U.S. workforce. The most valid predictor of job performance was work samples with a validity of .54. This is not surprising because work samples involve giving an applicant a test that resembles the job itself (e.g., giving a typing test to applicants for the job of typist).

A second source of literature for this paper includes evidence on the construct and criterion-related validity of situation judgment tests (SJTs). These tests have a long history and extensive literature. Several recent reviews have integrated much of this literature. McDaniel, Finnegan, Morgeson, Campion, & Braverman (2001) conducted a meta-analysis and found that SJTs predict job performance and are moderately correlated with  $g$ . McDaniel, Hartman, & Grubb (2003) studied the effect of different kinds of instructions on the validity of SJTs and found that instructions effected the correlations of SJTs with cognitive ability. McDaniel and Nguyen (2001) reviewed the literature on SJTs and drew conclusions about the construct validity of such measures.

### **3. Issues Addressed in this Critique**

This paper seeks to answer several issues concerning the measurement of tacit knowledge and the use of SJTs: a) we provide information about the similarities between Sternberg's tacit knowledge measures and commonly used SJTs in personnel selection; b) we question the existence of a general factor in SJTs; c) we describe the relationship between  $g$  and SJTs; d) we examine the extent to which SJTs are a method of measurement, rather than measuring a single construct; e) we compare the validity of SJTs with the validity of  $g$  for predicting job

performance; f) we provide suggestions for future research on tests of this type. Each of these issues is addressed in turn.

**a) Is Sternberg's description of tacit knowledge really the same as other well known and measured constructs?**

Although not acknowledged by Sternberg, situational items, similar to tacit knowledge items, have been in use for over 70 years. The earliest judgment test is the Social Situations subtest of the George Washington Social Intelligence Test (Moss, 1926), which was a failed attempt to develop a non-g “intelligence” test. During World War II, there were several applications of judgment tests in the U.S. Army (Northrup, 1989). Other examples of situational judgment tests include the Practical Intelligence Test (Cardall, 1942), the Business Intelligence Test (Bruce, 1965), the Supervisor Judgment Test (Greenberg, 1963), the Management Judgment Test used in the early 1950's and 1960's a part of the Early Identification of Management Potential program of Standard Oil (Campbell, Dunnette, Lawler, & Weick, 1970), and the U.S. Office of Personnel Management's Test 905 (Corts, 1980) of human relations skills in blue collar supervisors.

Motowidlo, Dunnette, & Carter (1990) examined the use of a situational judgment test, which they referred to as a low-fidelity simulation, for selecting entry-level managers. This measure, like other measures, presented applicants with verbal descriptions of work situations and several alternative courses of action for each situation. Wagner and Sternberg (1991) published a test called the Tacit Knowledge Inventory for Managers (TKIM). The purpose of this measure is to identify individuals whose “tacit knowledge” indicates the potential for successful performance in managerial or executive careers.

Although not labeled as a situational judgment test, the similarities between SJTs and the TKIM are clearly shown in the examples given below. The first item is an example of a situational judgment test in which a scenario is described and the respondent must identify an appropriate response from a list of alternatives or rate the effectiveness of each alternative. The item is from a World War II Army situation judgment test (Northrup, 1989).

A man on a very urgent mission during battle finds he must cross a stream about 40 feet wide. A blizzard has been blowing and the stream has frozen over. However, because of the snow, he does not know how thick the ice is. He sees two planks about 10 feet long near the point where he wishes to cross. He also knows where there is a bridge about 2 miles downstream. Under the circumstances, he should:

- A. Walk to the bridge and cross it
- B. Run rapidly across the ice
- C. Break a hole in the ice near the edge of the stream and see how deep the stream is
- D. Cross with the aid of the planks, pushing one ahead of the other and walking on them.
- E. Creep slowly across the ice.

The second item is an example item from the instructions for the TKIM provided in Appendix A of *Practical Intelligence in Everyday Life*.

Your immediate supervisor has asked you for your opinion on a new promotional campaign that she has developed. You think the promotional campaign is terrible and that using it would be a big mistake. You have noticed previously that your supervisor does not take criticism well, and you suspect she is looking more for reassurance than for an honest opinion.

Given the present situation, rate the quality of each of the following reactions on this 1- to 7-point scale.

1	2	3	4	5	6	7
extremely			neither good			extremely
bad			nor bad			good

- a. Tell her you think the campaign is great
- b. Tell her you like the work but have some reservations about whether it is the right campaign for this client

The actual items from the TKIM have more responses than are presented here. This example was taken from the instructions and is provided here for illustrative purposes. Each item describes a problem scenario that is followed by several potential responses to be evaluated by the respondent. Clearly, there is a great deal of similarity between the format of these two items.



McDaniel et al. (2001) compared the content of the items in the TKIM to other SJTs to demonstrate that the TKIM item content is not novel, unique, or a new concept in measurement. They note that both kinds of tests present situations that are hypothetical work-related scenarios that require either a judgment about the appropriateness of an action or a choice of the best action from among a range of options. McDaniel et al. (2001) describe two specific items in the TKIM that are very similar to items on the Supervisory Practices Test (Bruce, 1974) and the Teamwork-KSA Test (Stevens & Campion, 1999). One example is provided here. Item 4 of the TKIM asks the respondent to rate ten different strategies for handling the day-to-day work of a business manager. These strategies include thinking in terms of tasks instead of hours worked, being in charge of all phases of every task or project, spending time planning the best way to do a task, etc. Item 41 of the Supervisory Practices Test asks how the examinee would do daily work. Response options include taking care of the details while asking for planning from the boss, delegating details but planning with the boss, or handling details by oneself. Item 21 of the Teamwork-KSA Test asks the examinee to indicate which of four response options would most likely help the team to do its planning and coordinating. The options include examining past practices as a guide, considering priorities, pacing, sequencing of tasks and activities, etc. As these examples illustrate, there are obvious similarities between these purportedly different measures. This is not to say that the measures are identical, but that they measure similar content using similar methodology.

In sum, Sternberg's tacit knowledge item type is nothing new, but is simply one application of situational judgment items. Unfortunately, Sternberg and colleagues have ignored the situational judgment literature and this has given them an unsubstantiated optimism in finding a non-g-related measure of intelligence.

**b) Is there any evidence of a general factor in situational judgment tests?**

Sternberg et al. (2000, p. 223) stated that, “tacit knowledge appears to reflect a single underlying ability, which we label practical intelligence.” Gottfredson (2003) reviewed the limited evidence offered by Sternberg to support this claim and concludes that he offers “virtually no pertinent data” (p. 376). His claim can also be rebutted by several lines of evidence in the situational judgment literature. Northrop’s (1989) summary of U.S. Army research program concerning SJTs indicated that there is no general factor emerging from these tests. Northrop concluded that situational judgment does not form any clear factors and that such measures are *g*-saturated and multifaceted. Clause, Mullins, Nee, Pulakos, and Schmitt (1998) stated that in situational judgment tests, multidimensionality may occur within individual items and render the interpretation of factor structures hazardous.

We can illustrate the factorial complexity of SJT items by reviewing two items (see Figure 1). These items were developed for professional positions in a major corporation and are presented here with permission. Each item presents a scenario and several response options. The respondents were asked to rate the effectiveness of each response option for resolving the problem depicted in the scenario. Thus, each response option can be considered an item. Cognitive ability and personality test data were available for each respondent. In scenario A, all the displayed response options were correlated with both cognitive ability and agreeableness. Options 1, 2, 4 were judged effective by those high in cognitive ability and low in agreeableness. Option 3 was chosen by those who were low in cognitive ability and high in agreeableness. None of four effectiveness ratings was significantly correlated with conscientiousness or neuroticism. In scenario B, those who found response option 1 effective tended to be high in conscientiousness and low in agreeableness. Option 2’s effectiveness ratings were significantly

correlated with conscientiousness while option 3's effectiveness ratings were associated with cognitive ability. Those who found option 4 to be effective were high on cognitive ability and low on agreeableness. The response options for these scenarios, like most SJT items, are often construct heterogeneous at the item level. Factor analyses of such items do not yield clear factors. Tests made up of such items measure multiple constructs and have loadings on multiple factors. Given this situation, there is little possibility of any general factor emerging from such tests.

---

Insert Figure 1 about here

---

**c) Are SJTs correlated with *g*?**

Sternberg has asserted that there is a general factor of practical intelligence that is distinct from *g*. Sternberg et al. (2000) stated, "Thus, there is a growing body of evidence, obtained in work, school, and community settings, that suggests that tacit knowledge measures a construct distinct from academic intelligence" (p. 159). McDaniel et al. (2001) summarized data from 79 samples, including 16,994 individuals, and obtained a population correlation of .46 between *g* and SJTs (.36 uncorrected for measurement error).

Further, McDaniel et al. (2003) reanalyzed the McDaniel et al. (2001) data examining a response instruction moderator. They identified two categories of response to SJTs: behavioral tendency and knowledge. Behavioral tendency instructions ask examinees what they would do in a given situation, what they would most/least likely do in a given situation, or to rate/rank what they would most likely do. Conversely, knowledge tendency instructions ask examinees to select the best response, select the best/worst response, or to rate the effectiveness of various

responses. SJTs with behavioral tendency instructions correlated .23 (.18 uncorrected for measurement error) with *g*. SJTs with knowledge instructions correlated .55 (.43 uncorrected for measurement error) with *g*. Since neither .23 nor .55 is “generally uncorrelated,” we conclude that “practical intelligence” measures are correlated with *g*.

**d) Are SJTs a construct measure or a measurement method?**

Most of the debate concerning the constructs measured by situational judgment tests have an implicit assumption that there is a single situational judgment construct and the studies seek to understand this unitary construct. Several recent authors (Chan & Schmitt, 2002; McDaniel et al., 2001; Weekly & Jones, 1999) have argued that situational judgment tests are measurement methods. Like other measurement methods, such as employment interviews or job knowledge tests, situational judgment tests can be built to measure a variety of constructs. To assess interpersonal constructs, one can build tests or interviews with interpersonal items or questions. Alternatively, one can build tests in which cognitive ability or conscientiousness are the major determinants in answering the item correctly.

McDaniel et al. (2003) showed that the kind of instructions given for SJTs can change the constructs measured by the SJT. As mentioned above, agreeableness, conscientiousness, and emotional stability are all correlated with performance on SJTs. They found that the kinds of instructions provided can change the magnitude of the correlations, as shown in Figure 1.

---

Insert Table 1 about here

---

From these results, McDaniel et al. concluded that SJTs with knowledge instructions assess primarily *g* with some personality and that SJTs with behavioral tendency instructions

assess primarily personality with some *g*. In sum, these results show that SJTs are not a unidimensional construct measure, but a measurement method capable of measuring a wide variety of constructs.

**e) Is the validity of SJTs lower than the validity of *g* for predicting job performance?**

Sternberg et al. (2000) stated, “. . .we argue that practical intelligence is a construct that is distinct from general intelligence and that general intelligence is not even general but rather applies largely, although not exclusively, to academic kinds of tasks. Moreover, practical intelligence is at least as good a predictor of future success as is the academic form of intelligence that is commonly assessed by tests of so-called general intelligence” (p. xi-xii).

McDaniel et al. (2003) found that the criterion-related validity of SJTs is .32 for predicting job performance across 84 studies using 11,809 individuals. The validity was .32 for SJTs with behavioral tendency instructions and .28 for SJTs with knowledge instructions. The criterion-related validity of *g* tests for predicting job performance is .51 for jobs of medium complexity (Schmidt & Hunter, 1998). Although the validity for *g* is higher than for SJTs (.51 vs. .32), the correlations are not directly comparable because the validity of *g* has been corrected for range restriction and the validity of SJTs has not been similarly corrected. The range restriction correction for SJTs was not made because little is known about the range restriction of SJTs. Had such information been available, the validity of SJTs might be near the validity of *g*. However, the validity of SJTs for predicting job performance is NOT an endorsement of “practical intelligence” as a predictor of job performance. As shown above, SJTs are known to measure *g* and personality. Optimally weighted composites of *g* and personality typically exceed the validity of *g*. For example, a composite of *g* and conscientiousness usually yields a validity of .60 (Schmidt & Hunter, 1998).

**f) What are some suggestions for future research on tests of this type?**

Researchers should explore further the construct validity of SJTs and perhaps investigate constructs such as interpersonal skill (be nice to the boss) and specific knowledge (how to operate a machine), rather than making up “new” constructs (e.g., practical intelligence) when well-documented ones (e.g., *g* and personality) exist. We also should investigate ways to target situational measures to assess specific constructs. Researchers should examine the predictive validity of SJTs over *g* in predicting other important life outcomes, such as health, marital satisfaction, drug abuse, delinquency, and crime. Researchers also might examine whether SJTs may be superior to other assessment methods with respect to issues such as race, age, and sex differences, face validity, and resistance to applicant faking.

**4. Conclusion**

In Hans Christian Anderson’s folk tale titled “The Emperor’s New Clothes,” the emperor and his servants remarked how beautiful a set of clothes were that in fact did not exist. Sternberg and his colleagues have acted in a similar manner by advocating the merits of a construct that does not exist. A critical examination of research on practical intelligence shows that it lacks substance as a theory and as a construct. Specifically, there is no evidence of a general factor of practical intelligence. Further, there is substantial evidence that alleged measures of “practical intelligence” are, in fact, situational judgment tests that assess *g*, conscientiousness, agreeableness, and emotional stability.

## References

- Brody, N. (2003). What Sternberg should have concluded. *Intelligence*, 31, 339-342.
- Bruce, M.M. (1965). *Examiner's manual: Business Judgment Test*. Larchmont, NY: Author.
- Bruce, M.M. (1974). *Examiner's manual: Supervisory Practices Test* (Rev. ed.). Larchmont, NY: Author
- Campbell, J.P., Dunnette, M.D., Lawler, E.E., & Weick, K.E. (1970). *Managerial behavior, performance and effectiveness*. New York: McGraw-Hill.
- Cardall, A.J. (1942). *Preliminary manual for the Test of Practical Judgment*. Chicago: Science Research Associates.
- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cattell, R. (1987). *Intelligence: Its structure, growth and action*. New York: North Holland.
- Ceci, S.J. & Liker, J. (1986). Academic and nonacademic intelligence: An experimental separation. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical Intelligence: Nature and origins of competence in the everyday world* (pp, 119-142). New York: Cambridge University Press.
- Chan, D. & Schmitt, N. (2002). Situational judgment and job performance. *Human Performance*, 15, 233-254.
- Clause, C.S., Mullins, M.E., Nee, M.T., Pulakos, E.D., Schmitt, N. (1998). Parallel test form development: A procedure for alternative predictors and an example. *Personnel Psychology*, 51, 193-208.
- Corts, D.B. (1980). *Development and validation of a test for the ranking of applicants for promotion to first-line federal trades and labor supervisory positions*. Washington, DC: U. S. Office of Personnel Management, Personnel Research and Development Center, PRR-80-30. NTIS PB81137747.
- Gordon, R.A. (1997). Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence*, 24, 203-320.
- Gottfredson, L.S. (2003). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, 31, 343-397.
- Greenberg, S.H. (1963). *Supervisory Judgment Test Manual*. Technical Series No. 35. Personnel Measurement Research and Development Center. Bureau of Programs and Standards, Standards Division. Washington, DC: U.S. Civil Service Commission.

- Gustafsson, J.E. (1988). Hierarchical models of individual differences in cognitive abilities. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence, vol 4* (pp. 35-71). Hillsdale, NJ: Erlbaum.
- Horn, J.L. & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253-270.
- Horn, J.L. & Cattell, R.B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107-129.
- Hunter, J.E. & Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications.
- Lave, J. (1988). *Cognition in practice: Mind mathematics and culture in everyday life*. New York: Cambridge University Press.
- McDaniel, M.A., Hartman, N.S. & Grubb III, W.L. (2003, April). *Situational judgment tests, knowledge, behavioral tendency, and validity: A meta-analysis*. Paper presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology. Orlando.
- McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A., & Braverman, E.P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730-740.
- McDaniel, M.A. & Nguyen, N.T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment, 9*, 103-113.
- McDaniel, M.A., Schmidt, F.L. & Hunter, J.E. (1988). A meta-analysis of methods for rating training and experience in personnel selection. *Personnel Psychology, 41*, 283-314.
- Moss, F.A. (1926). Do you know how to get along with people? Why some people get ahead in the world while others do not. *Scientific American, 135*, 26-27.
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640-647.
- Northrop, L.C. (1989). *The psychometric history of selected ability constructs*. Washington, DC: United States Office of Personnel Management.
- Schmidt, F.L. & Hunter, J.E. (1993). Tacit knowledge, practical intelligence, general mental ability, and job knowledge. *Current Directions in Psychological Science, 2*, 8-9.
- Schmidt, F.L. & 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.



- Schon, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Sternberg, R.J. (2003). Our research program validating the triarchic theory of successful intelligence: Reply to Gottfredson. *Intelligence*, 31, 399-413.
- Sternberg, R.J., Forsythe, G.B., Hedlund, J., Horvath, J.A., Wagner, R.K., Williams, W.M., Snook, S.A., Grigorenko, E.L. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.
- Stevens, M.J., & Campion, M.A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management*, 25, 207-208.
- Wagner, R. K., & Sternberg, R.J. (1991). *Tacit Knowledge Inventory for Managers: User manual*. San Antonio, TX: The Psychological Corporation.
- Weekly, J.A. & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology*, 52, 679-700.

Figure 1

Two SJT Items and their Relationship to *g* and Personality

<b>Scenario A</b>		
	You assigned a very high profile project to one of your project managers. The project is very complex and involves the coordination of several other project managers. During each of the project update meetings, your project manager indicates that everything is going as scheduled. Now, one week before the project is due, your project manager informs you that the project is less than 50% complete.	
	Responses:	
1.	Personally take over the project and meet with the customer to determine critical requirements.	High <i>g</i> , low agreeableness
2.	Meet with the customer to extend the deadline. Talk with the project manager about how the lack of communication has jeopardized the company's relationship with the customer.	High <i>g</i> , low agreeableness
3.	Coach the project manager on how to handle the project more efficiently.	Low <i>g</i> , high agreeableness
4.	Do not assign any high profile jobs to this project manager in the future.	High <i>g</i> , low agreeableness

<b>Scenario B</b>		
	You lead a project that requires specific, accurate data to make business decisions. The data-capturing methods currently being used do not provide you with the information you need. Another department promised to provide you with the information, but failed to do so at the last minute. This setback delayed your project and you are certain that you still require the information to complete your project accurately.	
	Responses:	
1.	Do the time-consuming work yourself even though it is not technically your responsibility.	High conscientiousness, Low agreeableness
2.	Try to find another department that can provide you with the necessary information.	High conscientiousness
3.	Ask the customer for a deadline extension and explain that the other department failed to provide the necessary information.	High <i>g</i>
4.	Ask your manager to pressure the other department to deliver the information.	High <i>g</i> , low agreeableness

Table 1

Meta-analytic Results of Correlations Among Situational Judgment Tests with *g* and with Big 5 Measures

	<i>Knowledge Instructions</i>	<i>Behavioral Tendency Instructions</i>
<i>g</i>	.55	.23
Agreeableness	.20	.53
Conscientiousness	.33	.51
Emotional Stability	.11	.51

Adapted from McDaniel et al., 2003