ELSEVIER

Intelligence 30 (2002) 1–55

INTELLIGENCE

Dissecting practical intelligence theory Its claims and evidence

Linda S. Gottfredson*

School of Education, University of Delaware, Newark, DE 19716, USA

Received 24 February 2001; received in revised form 20 September 2001; accepted 8 November 2001

Abstract

Sternberg et al. [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] review the theoretical and empirical supports for their bold claim that there exists a general factor of practical intelligence that is distinct from "academic intelligence" (*g*) and which predicts future success as well as *g*, if not better. The evidence collapses, however, upon close examination. Their two key theoretical propositions are made plausible only by ignoring the considerable evidence contradicting them. Their six key empirical claims rest primarily on the illusion of evidence, which is enhanced by the selective reporting of results. Their small set of usually poorly documented studies on the correlates of tacit knowledge (the "important aspect of practical intelligence") in five occupations cannot, whatever the results, do what the work is said to have done—dethroned *g* as the only highly general mental ability or intelligence.

1. Introduction

Critics of the general intelligence factor, *g*, often assert that it is merely "book smarts" and, therefore, can provide little or no advantage in the real world. Among the various multiple intelligence theories (e.g., Gardner, 1983; Goleman, 1995; for critical reviews, see Davies, Stankov, & Roberts, 1998; Hunt, 2001; Lubinski & Benbow, 1995; Messick, 1992), Sternberg's triarchic theory of intelligence (Sternberg, 1985, 1988, 1997; Sternberg et al.,

* Tel.: +1-302-831-1650; fax: +1-302-831-6058.
E-mail address: gottfred@udel.edu (L.S. Gottfredson).

PII: S0160-2896(02)00085-5

2000) is the most explicit in positing separate intelligences for academic and practical affairs. State Sternberg et al. (2000, pp. xi–xii):

> [W]e argue that practical intelligence is a construct that is distinct from general intelligence and that. . . [it] 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 [*g*]. Arguably, practical intelligence is a better predictor of success.

This conclusion, they suggest (p. xii), is based on much evidence:

> [W]e have collected data testing our theories from many studies in many parts of the world with many different populations and have published most of these data (some are too recent to have been published) in refereed scientific journals.

Sternberg et al.'s (2000) claim is a bold and important one: bold because it seems to defy the huge edifice of research results showing that *g* forms the common backbone of all mental abilities; and important because, if true, it would require a major reorientation in scientific thinking on intelligence. Their summaries of the research can seem impressive at first glance, but the work itself has received little scrutiny from mainstream intelligence researchers. *g* theorists have criticized certain aspects of the work on practical intelligence (e.g., Barrett & Depinet, 1991; Jensen, 1993; Ree & Earles, 1993; Schmidt & Hunter, 1993), but, to my knowledge, only one (Brody, in press) has examined any part of it closely.[1]

My aim here is to provide a close and thorough examination of the theory and research that Sternberg et al. offer and how they offer it. I examine the concept of practical intelligence and then its supporting research. I look especially at the research on tacit knowledge, because Sternberg et al. (2000, p. xi) describe it as "one particularly important aspect" of practical intelligence and it is the one aspect that they measure. My examination is carried out against the backdrop of research on *g* and its real-world correlates (e.g., Brand, 1987; Gordon, 1997; Gottfredson, 1997, in press a, in press b; Jensen, 1998, Chaps. 9 and 14; Lubinski & Humphreys, 1997; Schmidt & Hunter, 1998). Brody (in press) has examined research with the Sternberg Triarchic Abilities Test (STAT) in educational settings. I, therefore, limit my scrutiny to tests of tacit knowledge, which have been used mostly in work settings.

I distill two theoretical propositions and six empirical claims from the latest accounting by Sternberg et al. (2000) of their work, *Practical Intelligence in Everyday Life*, that seem especially central to their case that practical intelligence is a general tool of equal or greater value than *g* in practical affairs. I have also consulted previous summaries of their work for this purpose (especially Sternberg, 1985, 1997; Sternberg & Wagner, 1993; Sternberg, Wagner, Williams, & Horvath, 1995; Wagner & Sternberg, 1986, 1990; Wagner, Sujan, Sujan, Rashotte, & Sternberg, 1999).

I quote extensively from key statements scattered throughout these publications for two reasons. First, despite their many publications on the subject, Sternberg et al. provide no single, clear, and full explication of their theory and research on practical intelligence to

[1] Others have examined triarchic theory in general (e.g., Kline, 1991, 1998; Messick, 1992), but not practical intelligence in particular.

which readers can turn. *Practical Intelligence in Everyday Life* (Sternberg et al., 2000) constitutes the most extensive accounting of their research program so far, but it provides more of a collage of related theorizing than a carefully developed model of practical intelligence.[2] And instead of collating into tables the data from two decades of research, the book gives the same unintegrated narrative summary of selected results, study by study, that has been published in similar form before (e.g., Sternberg & Wagner, 1993; Sternberg, Wagner, & Okagaki, 1993; Sternberg et al., 1995; Wagner & Sternberg, 1993). Second, readers can better assess the credibility of my conclusions if they hear from Sternberg et al. in their own words.

I also provide extensive tables of information. Although some may at first seem redundant with the text, they are essential for keeping track of the shifts in argument that Sternberg et al. have made over the years. Others are necessary for showing the full pattern of results that their body of research yields versus the pattern of results that Sternberg et al. (2000) report.

To preview my conclusions, Sternberg et al. (2000) fail to support their assertion that practical intelligence is not only distinct from academic intelligence (*g*) but also equals or exceeds *g* in its ability to predict everyday success. Sternberg et al. can support their two major theoretical propositions only by ignoring the most relevant evidence on *g* and making implausible claims about practical intelligence. As for their six empirical claims, none is supported by the evidence they offer. When their evidence is retrieved and examined closely, it actually contradicts two of the claims (empirical claims 1 and 3), illustrates the operation of *g* and not any new "practical intelligence" (claim 2), supports the claim only when interpreted in a heads-I-win-tails-you-lose manner (claim 4), fails even to address the claim (claim 5), and is seen to be greatly overstated for practical intelligence while systematically understated for *g* (claim 6).

2. Definition of practical intelligence

Sternberg et al. (2000, pp. 31, 97–98) describe practical intelligence as one of three "broad kinds of abilities" or "domains of mental processing" in Sternberg's (1985) triarchic theory of intelligence. As seen in Table 1, they are analytical (academic), creative, and practical. Although the relation is not entirely clear, the three abilities are said to "reflect" the three parts of triarchic theory, specifically, its componential, experiential, and contextual subtheories. As "broad abilities," analytical, creative, and practical skills seem to represent, respectively, analyzing information, generating ideas, and applying both to meet personal goals. When described as reflections of triarchic theory's three "domains of mental processing," they represent, respectively, the mental components that people use to process information, that they employ at various levels of experience on a task, and that they use in order to adapt to, shape, and select their environments.

[2] See also Rabbitt (1988, p. 178) on the triarchic theory being "more a comforting envelopment in jargon than a carefully thought-through functional model"; Kline (1991, 1998, pp. 141–142) on the theory's concepts being noncontingent (vacuous because not contingent on evidence) and "pseudoempirical"; and Messick (1992, pp. 377–380) on triarchic theory being more semantic than causal and more metaphorical than empirical.

In their more recent theorizing on intelligence as "developing expertise," Sternberg et al. have concentrated on the distinction between the first and third abilities, which they now refer to as intelligences and the first of which they now label, more restrictively, as "academic" rather than "analytical." Although the earlier triarchic theory seems to present the two

Table 1
Sternberg et al.'s (2000) definitions of academic (analytical) vs. practical intelligence

Three broad abilities[a]	Analytical intelligence	Creative intelligence	Practical intelligence
Three "broad abilities" ("process domains") that are "reflected" in Sternberg's (1985) triarchic theory of intelligence			
Ability to:[b]	solve problems	decide what problems to solve	make solutions effective
	learn from context and reason	cope with novelty	solve real-world, everyday problems
	think critically, analyze and evaluate ideas, solve problems, make decisions	go beyond what is given to generate novel and interesting ideas	implement ideas, the ability used when intelligence is applied to real-world contexts
Subtheory "reflected":[c]	componential (the components that people use to process information)	experiential (information-processing components are applied to tasks with which we have varying levels of experience)	contextual (information processing components are applied to experience in order to serve one of three functions in real-world contexts, which are adapting to, shaping, or selecting environments)
Relates intelligence to:[d]	internal world	experience	external world
STAT subtests:[e]	Analytical (verbal, quantitative, figural, essay)	Creative (verbal, quantitative, figural, essay)	Practical (verbal, quantitative, figural, essay)
As further elaborated in Sternberg et al.'s (2000) knowledge-based theory of practical intelligence			
Ability for:[f]	facile acquisition of formal academic knowledge		facile acquisition and use of tacit knowledge
Kind of knowledge:[g]	declarative (knowing that) inert		procedural (knowing how) action-oriented
Kinds of expertise:[h]	abstract, academic		practical, everyday
Value in real world:[i]	useful, important not very important		indispensible uniquely important
Measured by:[j]	conventional psychometric tests (e.g., IQ)		tacit knowledge tests

abilities somewhat as different stages in (or constraints on) the acquisition and concrete application of mental competencies, the newer theorizing tends to treat them as parallel capacities for acquiring different domains of knowledge. Thus, academic intelligence is said to be the "facile acquisition of formal academic knowledge," which is "declarative," "inert," and "abstract," whereas practical intelligence is the "facile acquisition and use of tacit knowledge," which is "procedural," "action-oriented," and "domain-specific" (see Table 1). In all their descriptions of the two abilities, however, Sternberg et al. place them on opposite ends of a continuum that ranges, on one end, from problem solving that is internal and abstract to that which, on the other end, is external and directly useful in the "real-world."

The following statements provide Sternberg et al.'s (2000) clearest definitions of practical intelligence.

1. "Practical intelligence is what most people call common sense. It is the ability to adapt to, shape, and select everyday environments" (p. xi).
2. "Adaptation, shaping, and selection [of environments] are functions of intelligent thought as it operates in context. It is through adaptation, shaping, and selection that the components of intelligence as employed at various levels of experience become actualized in the real world. This is the definition of practical intelligence used by Sternberg and his colleagues" (p. 97).
3. "Practical ability involves implementing ideas; it is the ability involved when intelligence is applied to real world contexts" (p. 31).
4. Referring in particular to the measurement of practical intelligence by the STAT, Sternberg et al. (pp. 97–98) state that its "practical questions address the ability to solve real-world, everyday problems."

Looking at the first two statements, it is not entirely clear how practical intelligence differs from Sternberg's (1997) more global "successful intelligence," which is an amalgam of all three intelligences (academic, creative, and practical):

> [A]lso termed the *triarchic theory*, successful intelligence is the ability to achieve success in life, given one's personal standards, within one's sociocultural context. Ability to achieve

Notes to Table 1:

[a] See Sternberg et al. (2000, pp. 31, 97).
[b] See Sternberg (1997, p. 47) and Sternberg et al. (2000, pp. 31, 97–98).
[c] See Sternberg et al. (2000, pp. 30–31, 97).
[d] See Sternberg et al. (2000, pp. 97–98).
[e] Sternberg Triarchic Abilities Test used in school settings (Sternberg et al., 2000, pp. 97–100).
[f] See Sternberg et al. (1995, p. 916).
[g] See Sternberg (1997, p. 11, 236) and Sternberg et al. (2000, p. 107).
[h] See Sternberg et al. (2000, p. 10).
[i] See Sternberg et al. (1995, p. 916), Sternberg (1997, pp. 11, 236), and Sternberg et al. (2000, p. 10).
[j] Sternberg et al. (2000, p. 144) rely on tests of tacit knowledge to measure practical intelligence in work settings.

> success depends on capitalizing on one's strengths and correcting or compensating for one's weaknesses through a balance of analytical, creative, and practical abilities *in order to adapt to, shape, and select environments*" (Sternberg et al., 2000, p. 93, first emphasis in original, second emphasis added).

The most crucial concept in practical intelligence theory is tacit knowledge. The emphasis on tacit knowledge stems from Sternberg et al.'s (2000, p. 103) "knowledge-based approach to measuring practical intelligence." Tacit knowledge, "as an aspect of practical intelligence, is experience-based knowledge relevant to solving practical problems." Tacit knowledge is the "important aspect" of practical intelligence because "much of the knowledge needed to succeed in real-world tasks is tacit," making it "an important factor underlying the successful performance of real-world tasks" (p. 104).

> In our work, we have studied many aspects of practical intelligence, although we have concentrated on one particularly important aspect of it, *tacit knowledge*, namely the procedural knowledge one learns in everyday life that usually is not taught and often is not even verbalized. Tacit knowledge includes things like knowing what to say to whom, knowing when to say it, and knowing how to say it for maximum effect. (Sternberg et al., 2000, p. xi, emphasis in original)

The three key features of tacit knowledge for Sternberg et al. are that it is (a) highly context-specific procedural knowledge, (b) that is acquired on one's own with little support from the social environment, and (c) is instrumental in attaining personal goals (Sternberg et al., 2000, p. 107). Sternberg et al. also describe it more colloquially as practical know-how and knowing the ropes. Sternberg (1997, pp. 236–237) gives a specific example, one which highlights well the personal expediency that tacit knowledge is often said to serve:

> Promotions are, in fact, a particularly good example of the importance of tacit knowledge to practical intelligence. The people who get promoted within an organization are usually the ones who have figured out how the system they are in really works, regardless of what anyone may say about how it is supposed to work. . . In many fields, what matters even more than the work you do is the reputation you build for that work, and reputation is not always tantamount to the quality of the work.

Accordingly, tacit knowledge is highly context-specific and goal-specific: "tacit knowledge is always wedded to particular uses in particular situations or in classes of situations" (Sternberg et al., 1995, p. 917; see also Sternberg et al., 2000, pp. 107–108). Sternberg et al. have, therefore, developed separate tacit knowledge tests for different job titles (life insurance salesperson, academic psychologist, business manager, Army platoon leader, and several others). These are the measures that they have "targeted specifically at practical intelligence" (Sternberg et al., 2000, p. 103).

Tacit knowledge tests pose from 7–19 problem-solving scenarios that incumbents have verified as important in their occupation (platoon leader and so on). Each scenario lists 6–16 potential actions to take, each of which respondents rate on a seven- or nine-point scale for either quality or importance (see Wagner, 1987, for examples of items on the academic psychology test and early versions of the management test; appendices in Sternberg et al., 2000, for copies of the sales [Tacit Knowledge in Sales] and most recent management test

[Tacit Knowledge in Management, TKIM]; and Hedlund et al., 1998, for the tests of military leadership [Tacit Knowledge in Military Leadership] at three levels.)

The tests have been scored in one of three ways, the first two using experts' typical responses as the standard and the third using accuracy of response (Sternberg et al., 1995, p. 918; see also Sternberg et al., 2000, p. 123): (a) giving points for answers that were more common among experts than novices (Wagner and Sternberg, 1985), (b) calculating squared deviations from the profile of answers obtained from a highly expert group (Hedlund et al., 1998; Wagner, 1987; Wagner & Sternberg, 1990; Williams and Sternberg, undated), and (c) summing responses to items that represent correct rather than incorrect or distorted application of rules of thumb (in sales; Wagner et al., 1999). Each test usually has several subscales. They have variously been tacit knowledge for (a) managing self, others, and career, (b) managing self, tasks, and others, (c) attaining global ("big picture") and local (immediate) objectives, (d) a combination of the latter two (e.g., global-self, global-task), and (e) attaining interpersonal and intrapersonal objectives.

3. The theoretical case for practical intelligence

Extensive empirical research has led many if not most intelligence experts to conclude that *g* is both a highly general mental ability and a relatively stable human trait. Many researchers, therefore, now consider *g* the core dimension of intellectual competence or their working definition of intelligence (see overviews by Carroll, 1993; Deary, 2000).

The *g* factor is not, of course, the only broad human ability. It is, rather, the *most general* ability. It seems, for this reason, to capture what most people mean by the term intelligence—a broad ability to learn and solve problems (to "catch on," "make sense of things," and "figure out" what to do). First discovered by Charles Spearman at the beginning of the 20th century, *g* has now been shown to exist—alone—at the apex of a hierarchy of mental abilities. The strata of the hierarchy are distinguished by the generality of the abilities at those levels, that is, by the range of tasks on which those abilities enhance performance.

Carroll (1993) provides the most exhaustive and definitive accounting of this *g*-capped hierarchy. Arraying abilities according to how specific vs. general they are, his "three-stratum" theoretical summary of the evidence assigns specific abilities to Stratum I and the most general to Stratum III. Placement was determined empirically by reanalyzing 450 previous data sets: Stratum II abilities represent the factors emerging from the common variance of the specific tests at Stratum I, and Stratum III abilities are the factors that emerge from the common variance of Stratum II abilities. Stratum I includes narrow abilities, such as spatial relations, spatial scanning, perceptual speed, associate memory, and free recall memory; Stratum II factors are the broad group factors, such as broad visual perception, general memory, and processing speed that suffuse the specific abilities in Stratum I; and Stratum III consists of *g*, which is the only factor that is common to all Stratum II factors (Carroll, 1993). In fact, *g* is the major component of all the moderately highly correlated Stratum II factors, which in turn are the major ingredients of the Stratum I abilities. Stratum II abilities, thus, consist mostly of *g* plus strong flavoring, so to speak, from independent sources of variance.

As Deary (2000, p. 11; see also Gustafsson, 1984) describes, the hierarchical, multiple-levels-of-generality model has unified models of intelligence that were once thought incompatible (e.g., Cattell, 1987; Spearman, 1927; Thurstone, 1938; Vernon, 1971). He refers to the model as a "semi-settled consensus" (p. 17).

The Cattell–Horn "Gf–Gc theory" of fluid and crystallized intelligence (Cattell, 1987; Horn & Cattell, 1966) is among those enfolded by the three-stratum hierarchy of mental ability. I shall say a bit about the Gf–Gc distinction because it figures prominently in later discussions. Intelligence researchers now accept the distinction between fluid intelligence (Gf) and crystallized intelligence (Gc). IQ test batteries, such as the Wechsler series, measure them both. Fluid intelligence refers to what might be called a person's mental horsepower, the ability to solve cognitive problems on the spot. Crystallized intelligence refers to very general mental skills (e.g., language) that have been developed—crystallized—from exercising fluid *g* in the past. Although not definitive, independent studies suggest that Gf is isomorphic with (correlates 1.0 with) *g* itself, or nearly so (Gustafsson, 1988) (hence, when I speak of *g* in this paper, I am, therefore, referring to fluid *g*.) These studies show that Gc correlates about .8 with *g*, which means that Gf and Gc are also correlated about .8 ($1.0 \times .8 = .8$). Carroll's (1993) massive reanalysis located fluid and crystallized intelligence in Stratum II of his scheme, but it yielded only one Stratum III ability—*g*.

Returning to the claims by Sternberg et al., it is precisely the intelligence experts' growing consensus about *g*'s generality and stability that Sternberg et al. must nullify in order to make their case that practical intelligence is coequal to *g*. Their theoretical case for practical intelligence, thus, involves an implicit two-part attack on *g*: (a) shrinking the apparent generality of *g* (by labeling it as only academic), so there is room to posit other intelligences that are crucial in other realms of life, and (b) shrinking *g*'s apparent causal power by arguing that it represents only a particular domain of knowledge, or learned expertise, rather than a stable, genetically rooted capacity (a trait) for learning and applying knowledge. We will see later how Sternberg et al. use their redefinition of *g* in terms of domain-specific knowledge to set up an empirical contest between practical intelligence (domain-specific tests of tacit knowledge) and academic intelligence (tests of *g*). Namely, can tests of tacit knowledge (each one of which is tailored to specific task domains in everyday life, such as life insurance sales) equal or exceed tests of *g* (which are tailored to no particular life domain) in predicting performance in the highly specific task domains targeted by the specific tacit knowledge test in question?

> *Theoretical Proposition 1:* g is not general; it seems so only because intelligence researchers have worn blinders. It is actually only a narrow academic ability, whereas everyday tasks require practical ability.

The case for practical intelligence begins with the argument that general intelligence is not general after all, despite evidence seemingly to the contrary.

> An enormous literature has emerged in the field of intelligence that is compatible with the notion that intelligence is a single entity, sometimes called *g*, or the general factor. . . (Brand, 1996; Carroll, 1993; Jensen, 1998). We challenge this view in the present book. In particular,

> 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... We believe that previous investigators have failed to find the importance of practical intelligence simply because they have never adequately measured it or, in most cases, made any attempt to measure it. By confining their efforts to a narrow band of tests, they failed to find a class of tests that would enhance not only their predictions but their theoretical models. (Sternberg et al., 2000, pp. xi–xii)

Or, as Sternberg states it more succinctly in his book on "successful" intelligence:

> This book has a very simple point. Almost everything you know about intelligence—the kind of intelligence psychologists have most often written about—deals with only a tiny and not very important part of a much broader and more complex intellectual spectrum. It deals with *inert intelligence*... [O]nce you expand the range of abilities that are measured, the general IQ factor disappears. (Sternberg, 1997, pp. 11–12, emphasis in original)[3]

Note that Sternberg et al. (2000) are actually making two separate claims here: (a) that there are other broad *intellectual* abilities ("intelligences") besides *g*, and (b) that *g*'s functional value in life is limited primarily to academic tasks. They explain away the contrary fact that "the scientific evidence in favor of what is called the *g* factor is overwhelming" (p. xii) by simply asserting that psychologists have not tried to measure anything else. In reality, many psychologists have worked hard and long over the decades—but in vain—to make the *g* factor disappear in a futile effort to develop useful mental ability tests that do not measure mostly *g* [e.g., see Humphreys' (1986) personal account and also Carroll's (1993) thorough review]. Sternberg (1985, pp. 7, 121–122) himself describes one particularly striking such effort—Guilford's unsuccessful attempt to validate his 150-factor "structure of intellect" model. In fact, Messick (1992, p. 382) describes how the major hierarchical theories of intelligence (Cattell, 1987; Vernon, 1971) reflect research on a considerably *broader* range of cognitive and conative traits than does Sternberg's triarchic theory.

Sternberg et al. (2000, p. 9) argue that "the alleged general factor of human intelligence" is not just narrow, but specifically academic. Appealing first to the reader's everyday observations, they suggest that the existence of separate academic and practical intelligences is obvious in our daily lives (p. 32):

> We see people who succeed in school and fail in work or who fail in school but succeed in work. We meet people with high scores on intelligence tests who seem inept in their social interactions. And we meet people with low test scores who can get along effectively with practically anyone. Laypersons have long recognized a distinction between academic intelligence (book smarts) and practical intelligence (street smarts or common sense).

They later provide specific examples of such disjunctions in apparent competence as evidence for separate practical and academic intelligences. As Hunt (1995, p. 105) sums up

[3] He also asserts later in the book (p. 94) that, with factor analysis, "you will always get a general factor, because it is in the nature of the statistical procedure." This is not true. Providing one concrete counterexample, the statistical procedure produces no general factor from personality tests (see Hogan, 1991, on the "big five" personality traits).

these sorts of anecdotes, "Accounts of low test scores who became Phi Beta Kappas or of high test scores who were incompetent workers are not germane to the issue at hand. The issue is how well the tests do on the average, not how well they perform in individual cases."

One need not posit a new intelligence to explain such disjunctions, of course. Differences in personality, motivation, and experience would suffice. I discuss this stratagem of argument by counterexample later, in conjunction with empirical claim 2, but will note here that such logic could just as easily be used to "refute" just about any important generalization in the social sciences, medicine, and other fields where causes typically are not *both* necessary and sufficient (fatty diets do not invariably cause heart disease or carcinogens cancer.) Such argument would also lead to an infinite regress of new, highly specific intelligences whenever an "intelligence" is less than perfectly predictive (say, of grades in different academic subjects or performance in different jobs). In other words, it would lead us straight to the bottom of Carroll's three-stratum model to highly specific Stratum I tests of narrow abilities or expertise. As we shall see, this describes well Sternberg et al.'s own tests of practical intelligence.

Turning to their nonanecdotal argument for distinct intelligences, Sternberg et al. (2000) suggest that different intelligences are relevant to different task domains. The major difference between their proposed academic and practical intelligences, they assert (pp. 32–34, emphasis in original), lies in the kinds of problem solving they facilitate:

> [The] difference is the sheer disparity in the kinds of problems one faces in academic versus practical situations. The problems faced in everyday life often have little relation to the knowledge or skills acquired through formal education or the abilities used in classroom activities. . . Everyone encounters problems to which solutions are neither readily available nor readily derivable from acquired knowledge. This type of problem solving, frequently experienced in daily life, is referred to as *practical problem solving*. . . The intellectual skills that individuals exhibit in finding solutions to practical problems may be referred to as *practical intellectual skills*. . . When combined, these skills are often referred to as practical intelligence.

Table 2 lists the attributes that Sternberg et al. associate, respectively, with academic and practical tasks. As indicated there, academic tasks are said to call for thought, not action; are imposed rather than chosen; are esoteric; and their answers and means of solution are highly circumscribed. In contrast, both the nature of the problem and the solution of practical tasks are said to be more ambiguous, and their solution (of which there may be several) requires

Table 2
Sternberg and Wagner's (1993) distinction between academic and practical tasks

"Academic" problems tend to:	"Practical" problems tend to:
(1) Be formulated by other people	(1) Require problem recognition and formulation
(2) Be well-defined	(2) Be ill-defined
(3) Be complete	(3) Require information seeking
(4) Possess only a single correct answer	(4) Possess multiple acceptable solutions
(5) Possess only a single method of obtaining the correct answer	(5) Allow multiple paths to solution
(6) Be disembedded from ordinary experience	(6) Be embedded in and require prior everyday experience
(7) Be of little or no intrinsic interest	(7) Require motivation and personal involvement

everyday experience and personal interest. The difference between academic and practical is, thus, a distinction between, on the one hand, the narrow, pedantic, disconnected theoretical and, on the other hand, the messy, meaningful reality in which people actually live. Both kinds of tasks are found throughout life, but "the proportion of problems that are practical rather than academic increases dramatically when one moves out of the classroom" (Wagner & Sternberg, 1990, p. 494).

This distinction between academic and practical tasks is critical to practical intelligence theory in several ways. It simultaneously allows Sternberg et al. (2000) to push *g* into a small academic box while opening a new and bigger box from which they can draw a practical intelligence. The distinction also sets the stage for their assertion, which they never test, that "intelligence as conventionally defined may be useful in everyday life, [but] practical intelligence is indispensable" (p. 1). And, very importantly, it also reflects the way they measure practical intelligence, which is to rely on tests of tacit knowledge whose items often conform to the practical attributes in Table 2.

It is an empirical question, of course, whether or not our mental and social worlds are divided into the two kingdoms they describe, one ruled by academic intelligence and the other by practical intelligence. We can ask, however, how much sense it makes even to suppose that task domains and, hence, corresponding "intelligences," would divide along the lines they suggest in Table 2. And why should we label one column "academic" and the other "practical"? Sternberg et al. (2000) do not explain. Why should IQ tests be consigned to the academic category? Sternberg et al. treat the decision as self-evident.

A moment's thought reveals that their distinction fails the reality test. Neither schools nor IQ tests limit themselves to posing tasks with mostly "academic" attributes, that is, clear-cut but esoteric problems, with all the necessary information, and with only one right method and one right answer. Academic subjects, such as history, composition, biology, literature, physics, and philosophy, when taught well, hardly model a regimented learning of settled questions and answers. Rather, good instruction poses tasks that often share many of the attributes of so-called practical tasks, such as requiring problem recognition and information seeking, having more than one means to a solution, and the like. By Sternberg et al.'s reasoning, IQ tests should predict school grades better than they do job performance, but they actually predict both about equally well (.4–.6; Hunt, 1995, p. 104).

As for IQ tests, many of them are essentially tacit knowledge tests. The very object of tests of crystallized intelligence, such as the Vocabulary and Comprehension subtests of the WISC, WAIS, and Stanford Binet IQ tests, is to assess the facility with which people have picked up information in everyday settings without direct instruction. That is the essence of tacit knowledge as Sternberg et al. define it. Most vocabulary, for instance, is tacit knowledge, complete with the difficulties of articulating it—explicitly defining words—when asked to do so (such difficulty, Sternberg et al. tell us, is characteristic of tacit knowledge). Sternberg (1985, p. 307) himself, in his book on triarchic theory; see also Sternberg, 1987; Sternberg & Powell, 1983), has argued similarly in a different context:

> [T]here is reason to believe that vocabulary is such a good measure of intelligence because it measures, albeit indirectly, children's ability to acquire information in context... Most

> vocabulary is learned in everyday contexts rather than through direct instruction... More intelligent people are better able to use surrounding context to figure out the words' meanings. With time, the better decontextualizers acquire the larger vocabularies. Because so much of one's learning (not just vocabulary) is contextually determined, the ability to use context to add to one's knowledge base is an important skill in intelligent behavior.

If some IQ tests are essentially tests of tacit knowledge, as Sternberg's assessment attests—and if being tacit is the measure of practical knowledge—then "conventional" tests of ability and aptitude cannot be cordoned off as academic.

There is, however, a telling difference between Sternberg et al.'s various tacit knowledge inventories, on the one hand, and, on the other, IQ tests that call for tacit knowledge, but it has nothing to do with the academic-practical distinction they propose. Specifically, the former are designed to assess highly domain-*specific* knowledge that few people may have had the opportunity to pick up (such as bank management) whereas IQ tests intentionally avoid such specificity. Rather, they are domain-*general*: they seek to assess broad cultural knowledge to which all individuals have been exposed ("why do we go to doctors?" or "what is the definition of 'sentence'?"). In short, neither schooling nor IQ tests can be squeezed into the "academic" column in Table 2, and the real distinction between tests of intelligence and tacit knowledge is the breadth vs. specificity of the competence they tap. A look at the four dimensions used for distinguishing aptitude from achievement tests (e.g., breadth of material sampled and tie to specific curriculum; Cleary, Humphreys, Kendrick, & Wesman, 1975; Lubinski & Dawis, 1992, p. 4) also suggests that IQ tests fall at one end of the specificity–generality continuum and tacit knowledge tests near the other.

Many tasks in everyday life likewise fail to respect Sternberg et al.'s academic-practical distinction because they exhibit mostly "academic" attributes. For instance, there are many problems in daily life that institutions and our compatriots impose on us (academic attribute 1 in Table 2), that have only one correct answer (academic attribute 4), or that require frankly academic skills, such as reading, writing, and arithmetic: filling out order forms, understanding instructions on prescription vials, using maps and bus schedules, calculating the amount of carpet needed for a room, understanding hospital consent forms, and comprehending instructions on preparing for an upper gastrointestinal tract radiographic procedure. These are but a few of the items from the National Adult Literacy Survey (NALS; Kirsch, Jungeblut, Jenkins, & Kolstad, 1993) and the Test of Functional Health Literacy in Adults (TOFHLA; Williams, Baker, Parker, & Nurss, 1998), two highly *g*-loaded tests representing everyday demands for self-care in modern life (Gottfredson, in press a, in press b). If such tasks are not highly practical for meeting one's personal goals, then the term has no meaning as Sternberg et al. use it.

As detailed further elsewhere (Gottfredson, 1997, in press a), *g* crosses the boundary between academic and practical, no matter how that boundary is defined. This cross-content generality of *g* is captured by Spearman's famous phrase, "the indifference of the indicator," which refers to the fact that *any* kind of test content or format (the indicator) can be used to measure the general factor, *g*, well. *g*'s effect sizes do range widely, but that variation has little or nothing to do with how intrinsically practical or personally consequential a task is. Rather, *g*'s utility rises when tasks are more complex, for example, when they are ambiguous,

unpredictable, evolving, multifaceted, lack complete information, or have unclear means–ends relations. Sternberg et al.'s (2000) research focuses on professional expertise in jobs, such as business manager and company commander, but the task demands that best distinguish complex, *g*-loaded jobs, such as these from simpler ones, are requirements for the very kinds of complex information processing that *g* exemplifies: for example, "deal with unexpected situations," "learn and recall job-related information," and "identify problem situations quickly" (Arvey, 1986, p. 418; Gottfredson, 1997, pp. 97–105). These requirements inherently involve ill-defined problems that require experience and may have many possible solutions, so Table 2 would seem to regard them as highly practical. That would make "practical" tasks, then, among the *most g*-loaded.

What five of the seven descriptors of "academic" tasks actually represent are rules for creating ability test items that will be *reliable* and *unbiased* and, thus, more *valid*. Test developers create items (academic attribute 1) that are well defined and have only a single correct answer (academic attributes 2 and 4) so that they will be more reliable. Although the accuracy of answers must be unambiguous, it matters not whether there are multiple ways to reach the answer (academic attribute 5 is not necessary). If the goal is to measure fluid *g* (mental "horsepower"), it is also important to provide all the necessary pieces of the puzzle to be solved (academic attribute 3) and not require any background information. If the goal is to measure crystallized *g* (general knowledge accumulated from using fluid *g* in the past), test items must avoid testing for information that is highly particular and, thus, not been available to everyone. Eliminating disparities in exposure is aided by disembedding the tasks from everyday experience (academic attribute 6). In short, because IQ tests are meant to measure a general capacity for solving problems of any type, they must avoid measuring the specialized knowledge necessary for learning and for solving some particular type—academic or otherwise—with which only a few have had experience. This also means that they *may* (not must) be of little intrinsic interest (academic attribute 7), as long as they are sufficiently engaging for individuals to try their best.

One difference between the tasks posed by IQ tests and by everyday life is, thus, the specificity of the skill or ability they measure best. As already noted, tests of aptitude and ability are designed as well as possible to exclude items that are sensitive to differences in exposure and experience, so they avoid items that tap knowledge for specific cultural or academic domains. In everyday life, however, people often differ enormously in the cultural domains they inhabit and the specific tasks they have undertaken and had a chance to master, so performance on everyday tasks—on life's specific "achievement tests"—reflects idiosyncratic exposure to a much greater degree than do IQ tests.

This raises the second difference between tests of IQ and tacit knowledge, which will also become very relevant when we consider the contest Sternberg et al. have set up between the two proposed intelligences. It is this. Although everyday life is often a highly *g*-loaded mental test, it is hardly a standardized one (Gordon, 1997; Gottfredson, in press b). As just intimated, we all take somewhat different life tests, so to speak, often limiting the range of task difficulty we choose to undertake. We can also call on other people's intelligence (get help) in performing life tasks that strain our capabilities. Such nonstandardization of the "test" items and "test taking" in daily life makes it more difficult to perceive *g*'s impact in everyday life,

because it requires careful effort to equate the "tests" and to isolate *g*'s effects from other factors known to influence performance, such as motivation, personality, experience, and special talents. As we shall see, Sternberg et al. (2000) capitalize on that nonstandardization to impute practical intelligence when other, uncontrolled differences among individuals and their circumstances could explain the phenomena they offer as evidence for a separate practical intelligence.

The criteria in Table 2 for defining academic tasks are, therefore, only matters of test format and manifest content. They confuse an explanation for *how* tests measure abilities well with *which* abilities they measure. They, therefore, fail to support Sternberg et al.'s (2000) theoretical argument for separate practical and academic intelligences.

Theoretical Proposition 2: g is not a trait, but situation-specific expertise. Practical intelligence, however, is both.

Sternberg et al.'s (2000) case against the generality of *g* takes a second form. If their first proposition limits the external reach of *g* to the domain of academic tasks, their second proposition restricts its internal depth to mere knowledge with only vague and tenuous biological roots. More specifically, Sternberg et al. try to create ontological parallelism for *g* and practical intelligence by arguing that, although *g* may have some limited generality, it is no deeper a trait than is practical intelligence.

The challenge in making this argument is that there is overwhelming evidence that differences in *g* represent a highly general and stable human trait, while there is none for practical intelligence. Sternberg et al., therefore, pursue a two-pronged strategy: to try to reduce IQ tests to the level of tacit knowledge tests (they measure only a specific kind of developing expertise) while they elevate tacit knowledge to the current status of IQ tests (they measure a general ability factor). That is, while empirical evidence accords *g* but not practical intelligence the status of a trait, Sternberg et al.'s theoretical argument does the opposite. The effort to, thus, turn the tables on *g* requires a convoluted series of incorrect assertions about *g* and inconsistent, implausible ones about practical intelligence.

The empirical evidence leaves no doubt that *g* is a trait and, specifically, that there is genetically rooted continuity in individual differences in *g* from infancy into old age. For instance, cognitive differences that are present in the first weeks of life correlate moderately well with childhood IQ; rank in childhood IQ changes little from year to year; and IQ becomes increasingly (and highly) heritable with age (80% by late adulthood). Evidence also shows that many of *g*'s biological, information processing, and socioeconomic correlates are not only heritable too, but that they also share some common genetic roots with *g* (e.g., Colombo, 1993; Jensen, 1998, Chap. 7, pp. 229–234; Lichtenstein & Pedersen, 1997; Moffitt, Caspi, Harkness, & Silva, 1993; Plomin & Bergman, 1991; Plomin, DeFries, McClearn, & McGuffin, 2000; Tambs, Sundet, Magnus, & Berg, 1989; Thompson, Detterman, & Plomin, 1991). Sternberg et al. (2000, p. 2) do not mention this evidence except to concede the bare minimum: their "view in no ways rules out the contribution of genetic factors" because "[m]any human attributes, including intelligence, reflect the covariation and interaction of genetic and environmental factors". All behavioral genetic knowledge

above that minimum, however, they implicitly and indirectly repudiate in order to argue that *g* is not a highly stable, strongly genetic trait.

Their effort to strip *g* of its status as a trait begins when they suggest that it is mostly a socially constructed phenomenon (just another form of "developing expertise") whose biological roots are at best thin and obscure.

> Some intelligence theorists point to the stability of the alleged general factor of human intelligence as evidence for the existence of some kind of stable and overriding structure of human intelligence. But the existence of a *g* factor may reflect little more than an interaction between whatever latent (and not directly measurable) abilities individuals may have and the kinds of expertise that are developed in school. With different forms of schooling, *g* could be made either stronger or weaker. In effect, Western and related forms of schooling may, in part, create the *g* phenomenon by providing a kind of schooling that teaches in conjunction the various kinds of skills measured by tests of intellectual abilities. (Sternberg et al., 2000, p. 9).

Nowhere do they discuss, let alone deny or explain, the evidence contradicting the statement they have just made—the evidence either for the relative stability of IQ over the lifetime, or for that stability originating in largely genetic factors, or for the emergence of virtually identical *g* factors in all age, sex, race, and national groups studied so far (Jensen, 1998, pp. 85–88; Plomin et al., 2000). Instead, they create the false impression that stability in age-normed mental competence (IQ) is a social accident rather than a biologically rooted fact when they assert, without evidence, that the *g* factor emerges because Western societies happen to teach together ("in conjunction") the separate skills which they then measure with IQ tests.

Sternberg et al. (2000, p. 1) explicitly reject the "conventional view of intelligence... [as] some relatively stable attribute of individuals" and propose, instead, the "alternative view... of intelligence as *developing expertise"* (p. 2, emphasis in original).

> [I]ntelligence tests [measure] an aspect, typically a limited aspect, of developing expertise... Developing expertise is defined here as the ongoing process of the acquisition and consolidation of a set of skills needed for a high level of mastery in one or more domains of life performance... Thus, conventional tests may unduly favor a small segment of the population by virtue of the narrow kind of developing expertise they measure. When one measures a broader range of developing expertise... [it] includes kinds of skills that will be important in the world of work and in the world of the family. (pp. 2, 9)

Sternberg et al. (2000) specifically reject the notion that there is an underlying general intelligence that causes differences in developed competence.

> We believe that the problem regarding the traditional model is not in its statement of a correlation between ability tests and other forms of achievement but in its proposal of a causal relation whereby the tests reflect a construct that is somehow causal of, rather than merely temporally antecedent to, later success. (Sternberg et al., 2000, p. 2)

They posit that test-outcome correlations result, not from enduring personal traits that affect subsequent behavior, but from both the antecedent and the consequent requiring overlapping knowledge ("developing expertise").

> According to this view, measures of intelligence should be correlated with later success, because both measures of intelligence and various measures of success require developing expertise of related types. (p. 2)

Sternberg et al. do point to common mental *processes* that affect the acquisition and use of different forms of expertise, but they describe ones that Sternberg (1985, pp. 338–341) has long presumed to be trainable and more like computer software than computer hardware, despite having acknowledged some genetic component.

> [P]erformance both on tests of intelligence and on indices of success typically require [sic] what Sternberg (1985) has referred to as *metacomponents* of thinking: recognition of problems, definition of problems, formulation of strategies to solve problems, representation of information, allocation of resources, and monitoring and evaluation of problem solutions. (Sternberg et al., 2000, p. 2, emphasis in original)

These are mental mechanisms that Sternberg (1985, p. 304) has described as being "centrally responsible for correlations between cognitive tasks and psychometric tests and for [whatever] limited success [that] psychometric tests [have] in predicting real-world performances of various kinds."

Sternberg et al. (2000) even downplay the notion that enduring individual differences in mental functioning of *any* sort might be consequential in everyday life when they suggest that personal attributes, whether malleable or not, play only a limited role in the development of intelligent behavior. The reason, as Sternberg (1985, p. 318) explains, is that intelligence must be traced to three loci: the individual, his or her behavior, and the contexts of behavior. Because "[i]ntelligence inheres in both the individual and the environments the individual inhabits," Sternberg believes it is "counterproductive to seek a unique locus of the nature of origins of intelligence when no single locus exists" (p. 318).

This view results in a contextualized, transactional definition of intelligence, where intelligence consists of intelligent (adaptive) behavior produced by a complex unit of which the person is only one component. Sternberg et al. (2000, p. 52) believe that:

> [The] individual and his or her context form a complex systemic unit [whereby] changes in the unit shape the content, dynamics, and adaptability of the individual's intellectual functioning in specific contexts.

The argument is, further, that intelligent behavior must be inferred from successful adaptation.

> [P]ractical intelligence... is defined as intelligence that serves to find a more optimal fit between the individual and the demands of the individual's environment, whether by adapting to the environment, changing (or shaping) the environment, or selecting a different environment. (Sternberg et al., 2000, p. 34)

This argument, it should be noted, shifts the criteria for defining practical intelligence from the objective task-based notion in Table 2 to a subjective outcomes-driven model in which intelligence seems to be whatever mental behavior helped the person adapt successfully. Adaptation itself is assessed against the person's own goals and particular circumstances, which renders the notion of intelligent behavior entirely relative.

> What this means is that there may be no one set of behaviors that is "intelligent" for everyone, in that people can adjust to their environments in different ways. (Sternberg, 1985, p. 310)

Accordingly, intelligence cannot be assessed in the same way for everyone. The components of mental hardware and software may be universal (Sternberg, 1985, pp. 52–53), but their development and expression is entirely relative to one's goals and subcultural context:

> No one or combination of the measurements [of intelligent behavior] would yield a definitive IQ, because any one instrument can work only for some of the people some of the time. Which instruments work for which people will be variable across people within and between sociocultural groups. (Sternberg, 1985, p. 312)

These theoretical assertions merely sidestep the pertinent empirical evidence that can expose them as false. In particular, the considerable behavioral genetic evidence for *g* and its correlates still sits in the wings ready to undermine the suggestion that *g* is socially constructed and not a strongly genetically rooted trait. Sternberg et al. (2000) nod sagely to behavioral genetics, but keep it off-stage by pointing to obvious but irrelevant truths. For instance, instead of learning some of the many relevant discoveries about the heritability (and joint heritability) of g, other abilities, achievement, and even our proximal environments (Plomin et al., 2000), we are told something obvious about what tests and behavioral genetics cannot do, namely, reveal what proportion of an *individual's* intelligence is genetic:

> Many human attributes, including intelligence, reflect the covariation and interaction of genetic and environmental factors. However, the contribution of genes to an *individual's* intelligence cannot be directly measured or even directly estimated; rather, what is measured is a portion of what is expressed, namely, manifestations of developing expertise. (Sternberg et al., 2000, p. 2, emphasis added)

No knowledgeable scientist argues, least of all behavioral geneticists, that the genetic component of an *individual's* IQ score can be estimated. That nonsensical question has never been the focus of heritability analyses. The aim of heritability analyses is quite different — to estimate what proportion of the phenotypic (observed) *differences* among us are the result our differences in genotypes, shared environments, and nonshared environments. For this purpose, behavioral geneticists have developed various ingenious methods for estimating the impact of these three sources of variation on phenotypic behavior. A naïve reader might be misled by Sternberg et al.'s reference to *individual* intelligence into supposing that the heritability of *differences* in intelligence cannot be estimated. They most certainly can — and have been.

Although Sternberg et al. concede that intelligence is somewhat genetically rooted, their discussion of intelligence as a "complex systemic unit" implies that the influence must be slight because the person is only one source ("locus") of that person's own intelligence. A reader would not guess that adult identical twins who were reared apart are almost as alike in IQ as are identical twins who grew up together (their IQs correlate .7–.8). Or that the heritability of IQ *rises* with increasing life experience, to .8 by old age (Plomin et al., 2000,

pp. 168–169). Sternberg (1997, p. 48) furthers the false impression that the genetic contribution to individual differences is unmeasurable and unstable when he states that:

> Intelligence is partially heritable and partially environmental, but it is extremely difficult to separate the two sources of variation, because they interact in many different ways. Trying to assign an average number to the heritability of intelligence is like talking about the average temperature in Minnesota. It can be as hot as the equator during the summer and cold as the North Pole during the winter. The heritability of intelligence varies depending on a number of factors.

Once again (except for the false analogy with Minnesota's weather), this statement is true but irrelevant.[4] The important point is not the truism that heritability (the ratio of genotypic to phenotypic variation) can vary, but that the variability in IQ heritabilities is patterned in theoretically important ways. Diversionary truisms like the foregoing one allow Sternberg et al. simultaneously to admit what cannot be denied (that individual differences in *g* and its correlates are genetically rooted and related) while denying its clear implication (that *g* is an enduring trait with causal power).

In their effort to strip trait status from *g*, Sternberg et al. (2000) have now brought us far afield from what is usually meant by an ability, let alone an intelligence. By their argument, it would seem that there can be no abilities, that is, tendencies to perform well on a broad class of tasks. This stance would be consistent, in fact, with their strategy for measuring practical intelligence using tests of tacit knowledge, which are targeted to "situation-specific" kinds of expertise whose development requires personal experience in relevant contexts.

The stance is not consistent, of course, with the triarchic theory's description of both academic and practical intelligence as "broad abilities" and "capacit[ies] to acquire" knowledge (see Table 1). Nor is it consistent with Sternberg et al.'s (2000) relentless effort to confer trait status on practical intelligence, which capacity is measured by tacit knowledge tests. While they are stripping *g* of its status as a trait, they are bestowing trait-like attributes on practical intelligence (i.e., tacit knowledge). In simply labeling practical abilities as an intelligence, they have instantly appropriated for "practical intelligence" all the connotations of generality and stability usually associated with IQ and *g*. Any inference of generality must be grounded in empirical evidence, of course, and evidence specifically that the *same* measured competence is useful—transferable—across *different* tasks. Sternberg et al. do offer evidence purporting to show the "domain generality" of practical intelligence (discussed as empirical claim 4 later), but it evaporates under the glare of independent inspection.

However, the dual claim itself—namely, that (1) IQ tests measure an expertise as domain-*specific* as do tacit knowledge tests but that (2) tacit knowledge tests measure a domain-*general* ability—might strike readers as a logical contradiction. Sternberg et al. (2000,

[4] Differences in IQ and other personal traits stem from differences in both our environments and our genes, and the heritability of such traits is calculated as the ratio of the genetic effects to genetic-plus environmental effects (i.e., the ratio of genotypic to phenotypic variability in intelligence). Were our environments to differ less over time, estimates of heritability would rise simply because our phenotypic differences (the denominator of the ratio) would shrink; conversely, were environments to become more different, the denominator would grow and resulting estimates of heritability fall.

p. 124), however, present it as a special achievement for tacit knowledge tests that conventional tests cannot claim:

> Tacit knowledge tests break down the artificial boundaries between achievement and ability testing... They are intended to measure both practical, experience-based knowledge and the underlying dispositions or abilities that support the acquisition and use of that knowledge.

In short, they suggest that tacit knowledge tests transcend the aptitude–achievement continuum while at the same time they shift IQ tests from the aptitude end of that continuum (where they belong) to its opposite pole, highly specialized achievement (where their tests of tacit knowledge belong).

To solidify their case against the general factor, *g*, Sternberg et al. supplement their theoretical arguments against it with pejorative labeling of *g* research and *g* researchers. In contrast to the "modern" ideas behind tacit knowledge tests, the research on *g* is "conventional" and motivated by researchers who, at best, cling to long-outdated notions and make patently silly "*g*-ocentric" claims (ones that they never actually do), such as that "overall performance from... employees... would be maximized" "if an employer were to use only intelligence tests" (Sternberg & Wagner, 1993, p. 1). And while told that Sternberg et al. "try to avoid contentious verbal arguments based on ideological position rather than scientific data" (Sternberg et al., 2000, p. xii), we elsewhere see pioneers Francis Galton and James McKeen Cattell ridiculed as the "public laughingstocks" that they should have been but were not in their time, the 19th century, for their forays into the psychophysical measurement of intelligence (Sternberg, 1997, pp. 54–55; but see Deary, 2000, Chap. 3 for an accurate history). The worldwide resurgence of research on speed of elementary cognitive processing (which has vindicated them) is dismissed scornfully. Mixing metaphors, Sternberg (1997, p. 55) contends that the resurgence is but a raising from the grave of a bad idea ("the bomb [that] proved to be a time-bomb") by "a crop of neo-Galtonians" who "have created a kind of night of the living dead" by "resurrect[ing] the work of Galton and Cattell."

Sternberg has even suggested that research on general intelligence is merely "quasiscientific" (Science and pseudoscience, 1999, p. 27). Whenever that research supports *g* theory, it may be telling us "less and less" (Sternberg, 2000, p. 372):

> General ability is not truly general, and its predictive value is more limited than it has seemed to be. Each study that suggests otherwise may be obfuscating rather than elucidating the nature of intelligence.

Thus, does he seem to condemn and dismiss the entire mainstream of research on intelligence.

4. The empirical case for practical intelligence

Sternberg et al. (2000) offer six kinds of evidence to support the validity of practical intelligence. The first five are meant to show that there exist separate practical and academic intelligences. The sixth is meant to show that tacit knowledge, the "particularly important aspect" of practical intelligence, predicts job performance at least as well as does *g*.

4.1. Empirical claim 1: laypeople distinguish between practical and academic intelligence

In an article summarizing the evidence on practical intelligence, Sternberg et al. (1995, p. 913) state that "[l]aypersons have long recognized a distinction between academic intelligence (book smarts) and practical intelligence (street smarts)." Their claim continues:

> This distinction... figures prominently in the implicit theories of intelligence held by both laypeople and researchers. Sternberg, Conway, Ketron, & Bernstein (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 analysis of the ratings supported a distinction between academic and practical aspects of intelligence for laypeople and experts alike. (Sternberg et al., 1995, p. 913)

Sternberg et al. (2000, p. 32) repeat this claim in their book, citing the same study: "This distinction is confirmed by research on the implicit theories of intelligence held by both laypersons and researchers."

Before examining their evidence, it is worth asking how pertinent such data might be for rendering judgments about the scientific merits of a theory. Sternberg et al. (1981) did, indeed, distinguish between the information value of *implicit* (or informal) and *explicit* (or formal) theories of intelligence. They described (pp. 38, 54) the former as "people's belief systems" and "word usage" that "serve as the basis of informal, everyday assessment... and training... of intelligence." That is, lay beliefs are important for sociological reasons, because they shape people's views of and, hence, behavior toward, themselves and one another. Sternberg et al. (1981, p. 37) described explicit theories as the "constructions of psychologists or other scientists that are based or at least tested on data collected from people performing tasks presumed to measure intelligent functioning" (e.g., "a battery of mental ability tests"). Implicit theories might, however, enhance the scientific study of intelligence if they "suggest aspects of intelligence behavior that... are overlooked in available explicit theories" (p. 38). In other words, lay theories are interesting but their value for scientific theories of intelligence is limited to hypothesis generation. Even if the claim were true, then, it would provide no evidence for the truth of any intelligence theory, including practical intelligence theory.

With that caveat in mind, let us nonetheless examine the claim and the evidence offered for it. Note first that the claim appeals partly to the very authority it is meant to repudiate intelligence experts: "laypeople and experts alike." Recall that Sternberg et al. (2000) began their book by arguing that mainstream intelligence experts are mistaken in their virtual consensus that *g* is general. Empirical claim 1, thus, appeals to intelligence experts' apparent good wisdom in agreeing with certain lay views of intelligence in 1981 as additional evidence against their supposedly misguided views today.

Where empirical claim 1 appears to give unquestioned credence to experts in 1981, Sternberg et al. seem to give them none today. Sternberg (2000, p. 365) now asserts that laypeople and experts have "a starkly different conception of intelligence" and that

laypersons' implicit theories about intelligence are *more* scientifically valid than the explicit (i.e., formal, evidence-based) theories of intelligence experts (Sternberg, 2000, p. 372):

> A case — I believe, a strong one — can be made that lay conceptions of intelligence better reflect the nature of intelligence than do the conceptions of many experts who are heavily involved in research on the phenomenon.

Sternberg never lays out his case, but he repeats the claim with equal certitude in a 1999 interview with *Psychology Today* (Epstein, 1999, p. 30):

> The professional concept of intelligence is much worse than the lay one. The problem is that many professionals have bought into the notion that intelligence is one single thing — an IQ, a *g*-factor. Our research pretty strongly shows that to be false.

Given this view, we might have expected Sternberg et al. (2000) to explain why experts seemed to hold views that supposedly supported separate intelligences in 1981 but not in 2000. Or why intelligence experts, in Sternberg et al.'s view, seem to have parted ways with both laypeople and the Sternberg Research Group, in the process veering away from the truth itself. However, Sternberg et al. (Sternberg et al., 1995, 2000) explain nothing. They simply point to the 1981 study without any comment, saying virtually nothing about it except that it supports their claim that "laypeople and experts alike" perceive separate academic and practical aspects to intelligence. Regarding Sternberg's (2000) claim that the two groups have "a starkly different conception of intelligence" today, he provides no support for his antiempiricist idea that we ought now to prefer lay views to scientific ones when seeking the truth about intelligence.

So what does the 1981 study actually show? I will go through it in some detail, not only to document how it repudiates the very claim for which it is invoked as support, but also to illustrate the manner in which Sternberg et al. tend to marshal evidence for practical intelligence theory.

The study consisted of three "experiments" (surveys), only the first two of which are relevant here. The first survey asked 186 laypeople in a train station, library, or supermarket to name behaviors that characterize one of three types of intelligence ("intelligence," "academic intelligence," and "everyday intelligence") or "unintelligence." Respondents listed 250 behaviors in all, 170 for the varieties of intelligence and 80 for "unintelligence." When asked to rate themselves on *all* three types of intelligence, the correlations among respondents' ratings were .80 (intelligence and academic intelligence), .60 (intelligence and everyday intelligence), and .44 (academic and everyday). Thus, despite the demand characteristics of this question (that there are, indeed, different intelligences), laypeople tended to rate themselves much the same on all three. Sternberg et al. (1981, pp. 41–42) concluded that "people seem to have at least somewhat different conceptions of the meanings of intelligence, academic intelligence, and everyday intelligence."

If we equate "everyday" with "practical" intelligence, then these "somewhat different conceptions" might seem to provide some support, albeit not strong, for Sternberg et al.'s (1995, p. 913) claim that "[l]aypersons have long recognized a distinction between academic intelligence (book smarts) and practical intelligence (street smarts)." But these are not the data to which they actually appeal as "support[for] a distinction between academic and

practical aspects of intelligence for laypeople and experts alike" (Sternberg et al., 1995, p. 913). Rather, they appeal to the results of a factor analysis conducted on ratings gathered in second survey, this one including intelligence experts as well as laypeople.

This second round of surveys took the list of behaviors produced in the first round of surveys, and asked the new respondents to rate each of the 170–250 behaviors, on a scale from 1–9, for either its "importance" (170 intelligent behaviors on Questionnaire 1) or "characteristicness" (250 intelligent or unintelligent behaviors on Questionnaire 2) for describing the "ideally intelligent" person. Respondents—both laypeople (recruited from the New Haven phone book) and intelligence experts—provided ratings of these many attributes for each of the three intelligences (their ideal concept of "intelligence," "academic intelligence," and "everyday intelligence"). Sternberg et al. (1981) then performed principal components analyses to extract independent factors from the "characteristicness" ratings (Questionnaire 2) for each of the three intelligences for both laypeople ($n = 28$, but see notes on Table 3 here) and experts ($n = 65$). Lay ratings for 98 behaviors and some unstated number of ratings from the experts were included. Except for two sets of factor loadings, all the results they reported are compiled in Table 3.

The study's authors (of which Sternberg was the principal one) concluded that the component factors of all three intelligences were "very similar" and shared a "common core" (Sternberg et al., 1981, pp. 50, 53). The common core also showed "remarkable similarities" when derived separately from lay and expert ratings (Sternberg et al., 1981, p. 46).

> [T]here seems to be a common core that is found in the belief systems of individuals in all of the groups we studied. The common core includes some kind of problem-solving factor, some kind of verbal-ability factor, and some kind of social-competence factor. (p. 53).

They then pointed out that the common core seen in these *implicit* theories shows up in experts' *explicit* theories.

> A recent review of literatures covering different approaches to understanding intelligence... concludes that these three aspects of intelligence plus a motivational one... seem to emerge from a variety of approaches to intelligence. (p. 53)

They next stressed the generality of this core:

> Thus, the results of the present research seem to converge with research of other kinds in suggesting that intelligence is found to comprise certain kinds of behaviors *almost without regard to the way in which it is studied*. These behaviors include (among possible others) problem solving, verbal facility, social competence, and, possibly, motivation. (pp. 53–54, emphasis added)

Because a social competence factor, not just strictly cognitive factors, also consistently emerged from the factor analyses, Sternberg et al. (p. 46) concluded that "the experts, like the laypersons, perceived intelligence as comprising quite a bit more than is presumably measured by IQ tests." Note, however, that this is not a practical intelligence factor of the sort that practical intelligence theory proposes, and it was usually the cognitive problem-solving factor to which they affixed the adjective "practical" (see Table 3)

Table 3
Factors obtained by Sternberg et al. (1981) from principal components analyses of about 100 traits[a] rated for their "characteristicness" of the "ideally _____ person" ("intelligent", "academically intelligent", "everyday intelligent"). Derived from ratings by laypeople vs. experts, and described by Sternberg et al. as representing either fluid ability (Gf) or crystallized ability (Gc)[b]

	"Intelligent"		"Academically intelligent"		"Everyday intelligent"	
Factors from ratings by laypeople and percent variance they explain (n = 28)[c]						
Gf	**Practical problem-solving ability**	**29**	**Problem-solving ability**	**8**	**Practical problem-solving ability**	**26**
Gc	**Verbal ability**	**10**	**Verbal ability**	**20**	Character	8
	Social competence	7	Social competence	7	Social competence	10
					Interest in learning and culture	6
Factors from ratings by experts and percent variance they explain (n = 65)[d]						
Gf	**Problem-solving ability**	**26**	**Problem-solving ability**	**26**	**Practical problem-solving ability**	**26**
Gc	**Verbal intelligence**	**23**	**Verbal ability**	**12**	Practical adaptive behavior	13
	Practical intelligence	9	Motivation	9	Social competence	16

[a] Sternberg et al. (1981, p. 44) report that 98 of the total 170 relevant lay ratings were factor analyzed. They report that only those ratings of traits that (a second set of) experts deemed most important were analyzed, but they do not report the number.

[b] Entries in bold represent the factors that Sternberg et al. (1981, pp. 46, 54) equated, as best I can discern, with either fluid intelligence (Gf) or crystallized intelligence (Gc).

[c] The *n* is inferred from Sternberg et al.'s (1981, p. 44) report that they used laypersons' results on Questionnaire 2 ($n = 28$) for this analysis.

[d] The *n* is inferred from Sternberg et al.'s (1981, p. 45) report that they used experts' "characteristicness" ratings. This would be Questionnaire 2 ($n = 65$), but there is some ambiguity because the "importance" ratings the analyses also relied on were from Questionnaire 1, which was administered to a different sample.

Finally, Sternberg et al. (1981) described the two most important factors explicitly in terms of *g*:

> Finally, the first two cognitive factors in the experts' conceptions of intelligence, like those in the laypersons' conceptions, seemed to correspond closely to fluid and crystallized abilities. (p. 46)

They amplified this point in the paper's concluding discussion:

> In particular, problem solving (or fluid ability) and verbal facility (or crystallized ability) seem to be integral aspects of intelligent functioning... In information-processing terms, crystallized ability seems best to separate the products of acquisition, retention, and transfer of verbal materials. These tests [of crystallized ability] primarily measure outcomes of previously executed cognitive processes rather than of current execution of these processes... Fluid ability tests, on the other hand, seem best to separate the execution of component processes of reasoning and problem solving and primarily measure current rather than past performance. (p. 54)

To summarize, the best stand-in for practical intelligence among the three a priori intelligences in the 1981 study is "everyday intelligence," but it is suffused with *g* by Sternberg et al.'s (1981) own account. On the other hand, respondents always viewed "intelligence" as highly correlated with "academic intelligence" and both as having some major "practical" component.[5] No matter which way the data are parsed, then, one can find a "practical" component, but it always comes in the company of Gf or Gc. Any disjunctions in perceived "intelligences" revealed by this study are like the differences among Stratum II factors in Carroll's scheme—they differ more in flavor than substance.

As described earlier, fluid *g* and crystallized *g* are both Stratum II factors in the hierarchical structure of mental abilities, they are highly intercorrelated, and fluid *g* seems isomorphic with the only higher-order Stratum III factor, *g*. This means that the 1981 study leads us, not to any new intelligence, but back to the old-*g*.[6] If it lends support to any theory, it is *g* theory, not practical intelligence theory.

Turning to the views of laypersons versus experts, Sternberg et al. (1981, p. 46) concluded the following:

> Thus, although there were differences between the exact factor structures obtained for laypersons and experts, the structures faithfully mirrored the high correlations between the two sets of ratings in indicating remarkable similarities in perceptions between people who professionally study intelligence and people who have no formal training in psychology, much less in the study of intelligence.

That is, it does not matter whether you ask laypeople or experts, or whether you ask them about intelligence, academic intelligence, or everyday intelligence, they always perceive the same set—"common core"—of competencies. Because fluid and crystallized *g* are the most important components of all three putative intelligences, all three are thereby suffused with the general ability factor, *g*. And their two biggest components are themselves both aspects of *g*. This is exactly the point that empirical claim 1 was meant to refute.

Authors have, of course, the prerogative to revise past conclusions in light of new knowledge, but Sternberg et al. never suggest any such reinterpretation. Rather, they routinely cite the 1981 study without comment as confirming their claims about lay theories of intelligence. They never mention the study's reliance on Gf–Gc theory. What are we to believe, then only that part of the study to which they vaguely refer us, but never specifically identify, that is said to show some sort of perceived distinction in forms of intelligent behavior? But which part is that the small distinctions that people perceive among the three a priori intelligences (intelligence, academic intelligence, and everyday intelligence), or the distinctions they perceive at a completely different level of analysis (namely, among the

[5] One cannot rely for clarification on Sternberg et al.'s application of the term *practical* and its frequent synonym *everyday*, because both are applied to so many and such different phenomena that they confuse as often as they clarify.

[6] Perhaps this is why Sternberg (2000, p. 365, emphasis added) would later assert, without explanation and without any hint of having reinterpreted the 1981 study, that "*none* [of these three components] correspond to a general factor and only the [second, verbal ability] corresponds well to abilities assessed by conventional intelligence tests."

component factors — problem-solving, verbal ability, and social competence — that they say constitute the *common* core of all three). Stated another way, is the putative academic-practical distinction revealed by looking across the rows in Table 3 or down its columns? The choice has very different implications for practical intelligence theory.

Do we also ignore the 1981 authors' conclusions on a related matter, specifically, the credence they gave to experts' theories at the time, theories that are actually much the same today but which Sternberg now describes as "strikingly different" from lay theories? More to the point, do we ignore the 1981 data suggesting that the implicit theories, lay or otherwise, are consistent with explicit theories of fluid and crystallized intelligence, that is, with Gf–Gc (and hence *g*) theory itself? In short, empirical claim 1 is credible only if we ignore the actual study that it cites.

Sternberg has moved away from *g*-based theorizing in the last 20 years, while more and more experts have moved toward it. If Sternberg et al. no longer stand by some of the 1981 conclusions, it would help readers to know which ones. However, any reinterpretation would have to be wholesale in order to support rather than undermine empirical claim 1.

4.2. Empirical claim 2: academic intelligence (g) cannot explain differences in practical problem solving, but the proposed practical intelligence probably does

Sternberg et al. have based this claim on the same few examples of problem solving each time they have summarized their evidence (Sternberg & Kaufman, 1998, pp. 494–495; Sternberg et al., 1993, p. 205; 1995, pp. 912, 915–916; 2000, pp. 34–38). This is how Sternberg et al. (2000, pp. 34–35, 38) describe the evidence:

> A number of studies have addressed the relation between practical and academic intelligence... Taken together, these studies show that ability measured in one setting (e.g., school) does not necessarily transfer to another setting (e.g., real-world task)... In other words, some people are able to solve concrete, ill-defined problems better than well-defined, abstract problems that have little relevance to their personal lives, and vice versa... What these studies... suggest is that there are other aspects of intelligence that may be independent of IQ and that are important to performance but have largely been neglected in the measurement of intelligence.

The claim rests on a handful of studies and two anecdotes of everyday activities where differences in performance seem to be independent of *g*. Most are cases of presumably low- to modest-IQ people being highly competent at some nonacademic task. The suggestion is that such examples contradict *g* theory and illustrate an independent practical intelligence at work. They fall into four categories:

1. *Individuals of presumably low IQ performed a task that seemed complex:* highly experienced but poorly educated milk processing plant workers found mental shortcuts that increased their efficiency in packing orders (Scribner, 1984, 1986); retarded children evaded elaborate security precautions to escape from a school for the mentally retarded (Sternberg et al., 1995, pp. 912–913); and Brazilian street children who did

badly on a formal math test nonetheless routinely performed mental math as street vendors (Carraher, Carraher, & Schliemann, 1985).

2. *Individuals of presumably low IQ performed a mental task that bright individuals could not:* highly experienced but poorly educated plant workers packed boxes of milk orders more efficiently than did their inexperienced white-collar substitutes (Scribner, 1984, 1986); and a much less taxing way to collect garbage, one that had not occurred to the PhD author, was instituted in the author's Florida neighborhood when a new, older worker was added to the work crew of mostly young high school dropouts (Sternberg et al., 1995, p. 912).
3. *IQ did not help predict who performed best in a particular nonacademic setting:* neither school grades nor test scores predicted which milk order packers were the best workers (Scribner, 1984, 1986); an arithmetic test did not predict differences in the frequency or correctness with which veteran supermarket shoppers in California used mental math when comparing products (Lave, Murtaugh, & de la Roche, 1984; Murtaugh, 1985); and the IQs among highly expert harness race handicappers did not correlate with their accuracy in predicting posttime odds (Ceci & Liker, 1986, 1988).
4. *IQ did not predict the complexity of the reasoning strategies people used to solve a problem:* solving the Sahara Problem (determining the number of camels that could be kept alive by a small oasis, Dörner & Kreuzig, 1983; Dörner, Kreuzig, Reither, & Staudel, 1983, articles in German cited by Sternberg et al., 2000, p. 37); managing a computer-simulated city (Dörner & Kreuzig, 1983; Dörner et al., 1983); and predicting posttime odds at the race track (Ceci & Liker, 1986, 1988).

All four categories represent the same strategy of arguing by counterexample. It is more a rhetorical device than a scientific strategy, however, because even high correlations between traits and outcomes, because they are less than 1.0, guarantee many exceptions to any general rule. We could just as well use such argument by counterexample to assert that smoking does not cause lung cancer. I might cite, for instance, an uncle who died of lung cancer without ever having smoked a single cigarette and a family friend who smoked heavily but remained cancer-free until her death at age 90. High intelligence may seldom if ever be a sufficient cause of life outcomes, but like smoking it certainly changes the *odds* of living a long, healthy, and productive life.

But let us return to the small collection of counter-examples offered. What does it illustrate? The examples represent people performing highly particular or atypical tasks, and seldom is enough information provided to determine what they illustrate about intelligence, if anything. The first three types actually appear to illustrate, not violations of *g* theory, but its very tenets. As described elsewhere (Gottfredson, in press a, in press b), *g*'s effects can vary widely across situations and groups, but they vary lawfully. For instance, *g* is a better predictor of job performance when tasks are more complex and when performers have more similar levels of experience and motivation. When differences in workers' experience are controlled, *g*'s predictive validities hold steady at successively higher average levels of job experience; when experience is not controlled, *g*'s effects are obscured and its validities are lower at successively lower average levels of experience (where differences in

experience are *relatively* greater and, therefore, have greater impact). Also, the greater the degree to which workers have been selected on *g* (that is, when there is more restriction in range on *g*), the larger any non-*g* factors will loom relative to *g* in explaining the workers' differences in performance. Differences in personality and motivation (in personnel selection parlance, the "will do" factors that affect job performance) can help predict performance, especially in simple tasks and more socioemotional ones. As with higher levels of *g* and motivation, longer experience and practice at a task (the "have done" factors) also enhance performance. Differences in relevant experience, however, tend to be most predictive where *g* is least predictive—when tasks are simple and workers differ considerably in task-specific experience. These well-documented regularities can explain the first three sets of examples.

As for the first form of putative evidence (i.e., dull people can do smart things), people of below-average IQ can successfully perform many specific tasks when they focus their practice on those tasks and when the tasks can be routinized, such as mentally totaling purchases while working as a street vendor. With keen motivation, dull individuals might even pool their information and experience to accomplish unexpected feats (a group of retarded children who individually failed even the easiest items on the Porteus Maze test nonetheless escaped from a secured facility).

Differences in motivation and relative experience probably explain most examples of the second type (dull people succeeded where smarter people failed). It should be no surprise, for instance, that an experienced, older garbage collector (of undetermined education and intelligence) working in Florida's summer heat and humidity might think of a faster way to do his job sooner than would the author sitting comfortably in his home. Nor should it be a surprise that highly experienced box packers outperformed their more educated but novice substitutes. With considerable experience, as military research has shown (Vineberg & Taylor, 1972, pp. 55–57; Wigdor & Green, 1991, pp. 163–164), low-ability workers can outperform inexperienced bright workers—although only until the latter get a bit of experience.

As for evidence of the third type (academic skills do not always predict differences in performance), all the examples are of narrow tasks performed by highly experienced people (box packers in a factory, veteran supermarket shoppers, long-time racetrack handicappers). None represents tasks that were novel to the individuals involved. Far from it, all were highly practiced. In addition, two of them were relatively simple (assembling milk orders, doing basic mental math). These represent precisely the sort of situation—highly practiced simple tasks—where *g* theory predicts that *g* will be relatively useless for forecasting differences in performance among *incumbents*. This does not imply that differences in mental ability are unimportant in training people for tasks that most people find very simple. For instance, the military services recruit nobody below the 16th percentile of mental ability and federal law forbids them to induct anyone below the 10th because of severe problems in trying to train and utilize low-ability recruits in years past, even for the simplest military jobs.

Regarding racetrack handicapping, the example hardly seems relevant. "These 30 men were highly experienced gamblers who, it turned out, had been attending races daily for 16 years, on the average" (Ceci & Liker, 1988, p. 96). Handicapping is also time-consuming: the men "typically devote six to eight hours handicapping ten eight-horse races" (Ceci & Liker, 1986, p. 132). These are men who were willing and able to devote most of their waking

hours to gambling: "they were able to afford to attend the races and bet nearly every day of their adult lives" (Ceci & Liker, 1988, p. 100). It, therefore, seems doubtful that any differences in these men's sophistication at a nonproductive endeavor would be explained by a new intelligence for dealing with the practical side of life.

The fourth kind of example (IQ does not predict the complexity of solutions offered) is not even relevant, because *g* predicts the *correctness*, not the *complexity*, of a solution. It is the complexity of a *task's* demands, not of the solutions people propose, that is core to *g* theory. Among the handicappers, the accuracy and complexity (completeness) of their implicit algorithms for predicting odds were correlated, but Rube Goldberg contraptions remind us that complexity and efficiency need not go hand in hand.

In short, none of these four kinds of evidence conflicts with *g* theory. None requires postulating a practical intelligence to explain the results. In no case was there evidence that the "practical" competence extended beyond the specific tasks in question, say, to health matters or even to everyday tasks of a similar nature. It is precisely such transferability or *cross*-task competence on a similar class of tasks, however, that is required to demonstrate a general ability.

Finally, it should be noted that the various tasks (e.g., packing orders, handicapping harness races, and solving the Sahara problem) that Sternberg et al. continue to cite constitute neither a large nor a meaningful sample of everyday tasks. Their more relevant examples (e.g., simple mental arithmetic in business encounters) tend to be simple, repetitive, and familiar tasks, so one need not posit any new intelligence to explain the success of even dull or poorly educated individuals in performing them.

4.3. Empirical claim 3: practical intelligence and academic intelligence have divergent developmental trajectories and, therefore, different etiologies

The claim is that practical and academic intelligences have "etiological independence (not necessarily complete)" because "the developmental trajectories of abilities used to solve strictly academic problems do not coincide with the trajectories of abilities used to solve problems of a practical nature" (Sternberg et al., 2000, p. 46). The claim is built from the well-known age trends in fluid and crystallized intelligence:

> *Fluid* abilities are those required to deal with novelty, as in the immediate testing situation... *Crystallized* abilities are based on acculturated knowledge... Using this distinction, many researchers have demonstrated that fluid abilities are relatively susceptible to age-related decline, whereas crystallized abilities are relatively resistant to aging... except near the end of life. (Sternberg et al., 2000, p. 39, emphasis in original; see also Sternberg et al., 1995, pp. 914–915)

The entire case for empirical claim 3 rests on equating practical with crystallized intelligence and academic with fluid intelligence. Sternberg et al.'s (1995, p. 914) theoretical rationale for this labeling is based on their task-based distinction between practical and academic intelligence as summarized earlier in Table 2.

> 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.

By the authors own definition of fluid intelligence (the ability to deal with novelty), however, one might have expected the opposite equation, namely, that ill-defined practical problems would require fluid intelligence and academic problems would require crystallized intelligence ("acculturated knowledge"). Recall, also, that Sternberg et al.'s (1981) study of implicit lay theories of intelligence had actually made the more expected equation, that is, matching fluid with *practical* intelligence:

> In particular, [practical] problem solving (or fluid ability) and verbal facility (or crystallized ability) seem to be integral aspects of intelligent functioning. (Sternberg et al., 1981, p. 54)

Although Sternberg et al. are not consistent in whether they associate practical intelligence with fluid or crystallized intelligence, it does not really matter empirically because the two are highly correlated, as noted earlier. Paradoxically, they are trying to forge a distinction between practical and academic intelligence by marrying it to the distinction between two highly correlated facets of *g*.

Moreover, *individual differences* in fluid and crystallized *g* are not etiologically independent, because the common variance of these highly heritable, highly correlated *g*'s — like other broad Stratum II abilities — seems to arise mostly from a common genetic substrate (Casto, DeFries, & Fulker, 1995; Jensen, 1998, pp. 122–126, 185–189; Plomin & DeFries, 1998). By tying their distinction between academic and practical intelligence to that between fluid and crystallized intelligences, Sternberg et al. (2000) effectively repudiate their own case for the etiological independence of their two proposed intelligences. Once again, the evidence they offer, when examined closely, proves the opposite of what they claim.

Ignoring this complication (the unmentioned close correlation between individual differences in fluid *g* and individual differences in crystallized *g*), Sternberg et al. (2000) point instead to less relevant data to support their claim: age trends in average scores from early to late adulthood. They draw attention, in particular, to the falling averages for fluid *g* but the steady or rising averages for crystallized *g*. They begin their argument by stating: "In particular, the idea that practical and academic abilities might have different developmental trajectories was supported in a number of studies" (p. 40).

They then cite several studies that measured everyday problem solving in addition to performance on "traditional" cognitive tests. Referring to the first (Denney and Palmer, 1981):

> [Performance on] traditional analytical reasoning problems (e.g., a "20 questions" task)... declined almost linearly from age 20, onward... [but performance on] problem solving task[s] involving real-life situations (e.g., "If you were traveling by car and got stranded out on an interstate highway during a blizzard, what would you do?")... increased to a peak in the 40- and 50-year-old groups, declining thereafter. (Sternberg et al., 2000, pp. 40–41)

Sternberg et al. use a second study (Williams, Denney, & Shadler, 1983) to justify their labeling the first ability (“analytical reasoning”) as academic and the latter (“problem solving [in] real-life”) as practical. They point, in particular, to how older adults had explained their continued everyday competence despite waning mental abilities: most of them thought that their “ability to think, reason, and solve problems had actually increased over the years,” despite evidence to the contrary on traditional tests, because they were referring to “solving kinds of problems different from those found on psychometric tests... [and which are] of an everyday or financial nature” (Sternberg et al., 2000, p. 40). As Sternberg et al. themselves point out, however:

> The available evidence suggests that older individuals compensate for declining fluid abilities by restricting their domains of activity to those they know well... and by applying specialized procedural and declarative knowledge. (Sternberg et al., 1995, p. 915; see also Sternberg et al., 2000, p. 42)

That is, they rely on past expertise rather than developing new forms of it, which hardly implicates the operation of some distinct practical intelligence. Indeed, a look at the cited study (Williams et al., 1983) shows that the elderly respondents reported being more afraid of making mistakes than when they were younger; having fewer and easier problems to solve than do younger people; and being better now at solving problems because they have more experience, are less emotional, and can take more time. None of this reflects a new and distinct intelligence, but only ways to compensate for general intellectual decline.

A third cited study (Cornelius & Caspi, 1987) provides more direct evidence on empirical claim 3 because it specifically measured both fluid *g* (completing a letter series) and crystallized *g* (verbal meanings) as well as everyday problem solving (e.g., dealing with a landlord who won’t make repairs, filling out a complicated form). Cross-sectional age trends in averages for the two *g*’s showed their typical divergence in adult development, with the trend for everyday problem solving being more similar to the one for crystallized *g*. These are, as Sternberg et al. (2000, p. 41) say, “similar results” to the others just mentioned. However, this third study revealed an awkward consequence of the empirical fact they continued to ignore: namely, because the two *g*’s are highly correlated, individual differences in everyday problem solving were found, not surprisingly, to be *equally* correlated with crystallized and fluid *g* (.27 and .29). If everyday problem solving is supposed to reflect crystallized intelligence (which they had designated as “practical”) and *not* fluid intelligence (designated “academic”), the former correlation should have been notably higher than the latter. Ignoring this obvious contradiction of their assertion that everyday problem solving reflects practical (crystallized) *rather than* academic (fluid) ability, Sternberg et al. (1995) simply create the impression that the fit between everyday intelligence and crystallized *g* may not be a snug one. Echoing Cornelius and Caspi (1987, p. 915), they state that “despite their similar developmental functions,” everyday problem solving among adults is “not reducible to crystallized ability,” presumably because the correlation is modest (.27).

But another study (Willis & Schaie, 1986) causes even worse complications for Sternberg et al. precisely because it does indeed find a snug fit for everyday problem solving, but with *both* crystallized *g* (.78) and fluid *g* (.83). This pair of very high correlations suggests that differences

in everyday problem solving might conform closely to both *g*'s, meaning that a single general factor might run through all forms of problem solving. However, Sternberg et al. mention this fact only to dismiss its clear relevance. When responding to Barrett and Depinet's (1991) conclusion that the Willis and Schaie (1986) study demonstrated that an "extremely high relationship existed between intelligence and performance on real-life tasks," Sternberg et al. (1995, p. 924) rejected that conclusion because the study's measure of everyday problem solving was, in their view, more academic than practical: it was "a paper-and-pencil psychometric test" of everyday skills that were "decidedly more academic than changing a flat tire or convincing your superiors to spend a million dollars on your idea."

The test in question, the ETS Basic Skills Test, required reading paragraphs, letters, guarantees, maps, and charts, as does the NALS mentioned earlier. Sternberg et al. (1995) do not explain why such skills are not practical ones. The implication seems to be that they are academic simply because they require reading, although that skill is one of the most essential in modern life: people with weak "functional literacy" skills "are not likely to be able to perform the range of complex literacy tasks that. . . [are] important for competing successfully in a global economy and exercising fully the rights and responsibilities of citizenship" (Baldwin, Kirsch, Rock & Yamamoto, 1995, p. 16). Nor do Sternberg et al. explain why such supposedly academic skills would correlate very highly with both crystallized and fluid *g* if the latter two really do reflect separate practical and academic intelligences. In their book, Sternberg et al. (2000, p. 39) describe the rejected Willis and Schaie (1986) study in another context, but immediately imply that it was problematic because it was just cross-sectional, although the same complaint would apply to the studies they themselves cite a page later to support their divergent etiologies claim. It might also be noted that their own tacit knowledge tests are "paper-and-pencil."

As with the prior empirical claim, Sternberg et al.'s (2000) evidence for this one is more consistent with *g* theory than practical intelligence theory. It is made to appear supportive only by highlighting marginally relevant data while ignoring or belittling directly relevant results that contradict the theory.

4.4. Empirical claim 4: tacit knowledge tests measure a general factor of practical intelligence

The final conclusion that Sternberg et al. (2000, p. 223) draw from their program of research on practical intelligence is that "tacit knowledge appears to reflect a single underlying ability, which we label practical intelligence."

> Although the kinds of informal procedural knowledge measured by tacit knowledge tests do not correlate with traditional psychometric intelligence, tacit knowledge test scores do correlate across domains. Furthermore, the structure of tacit knowledge appears to be represented best by a single general factor. (p. 159)

Recall that, although Sternberg et al. (2000) dispute any claim that *g* represents a truly general intelligence, they do accept the evidence that it is general within the realm they have labeled academic, which includes "conventional" mental tests. That psychometric generality,

as limited as they view it, was established empirically via factor analyses of many batteries of diverse tests, some in representative samples of the population (Carroll, 1993). What is the analogous evidence for a general factor of practical intelligence, specifically, "[t]he ability or propensity to acquire tacit knowledge... that conventional ability tests do not adequately measure" (Sternberg et al., 2000, p. 111)?

Sternberg et al. (2000) offer two kinds of evidence. The first is that different parts of the same tacit knowledge test measure a common factor. The second is that different tacit knowledge tests correlate with each other. To the extent that test parts or wholes intercorrelate and measure a common factor, Sternberg et al. describe this commonality as evidence for the "domain *generality*" of the *ability* measured by tacit knowledge tests. To the extent that they fail to correlate or measure a common factor, the results are interpreted as evidence for the "domain *specificity*" of the *knowledge* measured by tacit knowledge tests. Either way the results turn out, in other words, Sternberg et al. offer them as evidence for the theory; they provide either "convergent validity" or "discriminant validity." Such a heads-I-win-tails-you-lose procedure is incapable of falsifying any hypothesis.

Illustrating the first type of evidence offered, a study of 91 psychologists and 64 managers showed, via principal components analysis of each test's six component scales (self-local, task-local, etc.), that the job-specific tacit knowledge test given in each sample (one on psychology and one on management) was mostly unidimensional (Sternberg et al., 2000, p. 159; Wagner, 1987, pp. 1242, 1244–1245). Recall that, with only one exception (the sales test), responses to tacit knowledge tests are scored not for their accuracy but for their similarity to experts' responses. Sternberg et al. (2000) implicitly offer the foregoing two analyses as analogous to the factor analyses of the subtests of the major IQ test batteries, which typically have about a dozen subtests and always score responses for their accuracy. Even if granted the tenuous analogy, the separate factor analyses of the psychology and management tests cannot support the pertinent point, namely, that the two tests measured the *same* general factor, which is what Sternberg et al.'s labeling implies. To wit, Sternberg et al. (1995, pp. 919–920) summarized the separate analyses as both showing the "domain generality" of tacit knowledge. The claim is repeated in the section of their book entitled "Tacit Knowledge as a General Construct" (Sternberg et al., 2000, p. 159): for managers, the "analyses suggested a general factor of tacit knowledge," and for psychologists, "[a]s with the study of managers, the factor analytic results suggested a single factor for tacit knowledge within the domain of academic psychology." The apparent unidimensionality of the two individual tests provides no evidence, however, that these tests with markedly different content, given to separate samples, both measure the same common factor, but labeling each as "domain general" can create the illusion of evidence.

Sternberg et al. (Hedlund et al., 1998) also explored the factor structures of two military leadership tests, but for a different reason—to increase the tests' poor prediction of leadership performance. That is, they undertook the factor analyses of the questions for platoon leaders ($n = 368$) and company commanders ($n = 163$) not to assess the dimensionality of the tests, but to create more predictors by ferreting out any *multi*dimensionality in their tests. The original test scores seldom predicted any of the performance ratings above chance levels, and they wanted to "identify potential subsets of items that may provide additional prediction of

leadership effectiveness'' (p. 198). The search was somewhat successful at the company commander level: a five-question factor and a seven-question factor each correlated significantly with one of the nine one-item performance ratings[7] (Hedlund et al., 1998, pp. 29–30). Sternberg et al. (2000) fail to note that this success might demonstrate a lack of the ''domain generality'' they had earlier pointed to as important in the study of 91 psychologists and 64 managers. Nor do they report one case of complete lack of ''domain generality'' in a test—specifically, that the ''local'' and ''global'' halves of their test of TKIS never correlated above chance levels in any of the four samples to which they administered the test (two each of salespeople and undergraduates). As the original study had found, the correlations ''were not reliably different from 0'' (Wagner et al., 1999, pp. 163, 165).

Sternberg et al.'s (2000) second and more pertinent kind of evidence for a general construct of practical intelligence comes from four samples where the same respondents took two different tacit knowledge tests. The evidence is inconsistent, however, and, once again, so too are their conclusions. In a sample of 60 Yale undergraduates with no experience in either psychology or management, tacit knowledge for the former correlated .58 with tacit knowledge for the latter. Sternberg et al. (2000, p. 159) conclude from this correlation that ''individual differences in tacit knowledge are generalizable across domains.'' They later report for the military study, however, that two forms of tacit knowledge that one might have thought to be more similar—tacit knowledge for management and for military leadership—yielded lower correlations: .36 for platoon leaders, .32 for company commanders, and −.06 for battalion commanders (first row in bottom panel of Table 6). Sternberg et al. (p. 197) concede that ''the magnitude of this correlation does not indicate that the [two tacit knowledge tests] are measuring the same construct,'' but they suggest nonetheless that it ''may... reflect an underlying ability to acquire and use tacit knowledge that generalizes across performance domains, which is considered an important aspect of practical intelligence.''

Later, when they examine these two tacit knowledge tests' ability to predict leadership ratings, they are pleased that tacit knowledge for leadership produced a small increase in variance explained above and beyond that provided by tacit knowledge for management, because they suggest that this increase ''provides further support for the domain *specificity* of tacit knowledge'' (p. 203, emphasis added). The finding of support (either domain specificity or domain generality of the tests) from inconsistent evidence conforms to the fundamental inconsistency within the theory itself, which argues that tests of tacit knowledge measure *both* domain-specific knowledge *and* a domain-general tendency to acquire tacit knowledge of any type.

The claim that tacit knowledge represents a general factor of practical intelligence, however, is the very crux of the contest that Sternberg et al. have set up with *g*. As just

[7] There were 20 tacit knowledge questions on the test for company commanders, each question having from 4 to 16 possible answers, all of which respondents were asked to rate from ''extremely bad'' to ''extremely good.'' Scores were calculated as squared deviations from a profile of experts' responses, and then adjusted for level of disagreement among experts' responses on each option and for each soldier's tendency to use the whole rating scale (Hedlund et al., 1998, pp. 12–14).

seen, there is scant support for a claim that tacit knowledge reflects a general ability (of any sort), partly because there are virtually no pertinent data. As noted, only two studies, one of Yale undergraduates and one of Army officers, measured two forms of tacit knowledge using the same subjects. Ideally, one would want to factor analyze batteries of such tests in a wide range of populations — as has been done to verify the stability and generality of the general intelligence factor, *g*. And one would want to be able to rule out *g* as a potential source of correlation between tacit knowledge tests.

4.5. Empirical claim 5: practical intelligence is independent of academic intelligence (IQ)

Sternberg et al. (2000, p. 159) claim not only that tacit knowledge reflects a single underlying ability, but that it measures one that is "distinct from general academic intelligence." As direct evidence of this, they point to insignificant correlations between IQ scores and tests of tacit knowledge.

> Tacit knowledge is not a proxy for general intelligence... In study after study, this important aspect of practical intelligence [tacit knowledge] has been found generally to be uncorrelated with academic intelligence as measured by conventional tests in a variety of populations and occupations and at a variety of age levels. (pp. 111, 144)

The "variety of populations and occupations and... age levels" to which they refer is listed in Table 4: four samples of inexperienced college undergraduates, one of inexperienced Air Force trainees, one of civilian workers (experienced managers), and three of Army officers (experienced platoon leaders, company commanders, and battalion commanders). The 13 correlations in bold are those reported in Sternberg et al. (2000), and the remaining 14 were obtained from earlier published (Wagner, 1987; Wagner et al., 1999) and unpublished reports (Hedlund et al., 1998). Of the 27 correlations with IQ, only seven are significant. Weighted by sample size, the average correlations are .17 for the undergraduates, .07 (with the four ASVAB composites, not an IQ test) for the Air Force trainees, .14 for the managers in leadership training, and .13 and .12 for the Army officers, respectively, on two measures of tacit knowledge, one targeted to the officers' jobs (TKML) and one not (TKIM).

While the correlations are small, Sternberg et al. tend to overstate what they refer to as their "trivial[ity]" in civilian samples. First, they misstate the data. Sternberg et al. (2000, p. 157) report that "[i]n all the above [civilian] studies..., tacit knowledge test scores correlated insignificantly with *g*," but Table 4 shows that three of the seven civilian correlations were statistically significant. Sternberg et al. had not reported these three quantities (.30, .40, .25), but had, however, specifically said that the first was *not* significant (p. 147) and left the clear impression that the other two were not either (p. 151).[8] In a different chapter, they do mention

[8] In the first instance, "Again, the tacit knowledge scores did not correlate with verbal reasoning ability" (Sternberg et al., 2000, p. 147); in the second instance, "The total scores for undergraduates were uncorrelated with verbal reasoning scores" (p. 151); in the third, "tacit knowledge scores again did not correlate significantly with verbal reasoning scores" (p. 151).

Table 4 t4.1
Correlations between tests of tacit knowledge and IQ in nine samples t4.2

Tacit knowledge test	IQ Test	Sample	*n*	Correlation	
Academic psychology	DAT Verbal Reasoning	Yale undergraduates	22[a]	**−.04**	t4.4
	DAT Verbal Reasoning	Yale undergraduates	60[b]	.30*	t4.5
Management	DAT Verbal Reasoning	Yale undergraduates	22[c]	**.16**	
(various successive	DAT Verbal Reasoning	Yale undergraduates	60[d]	**.12**	t4.6
versions of TKIM)	ASVAB composites	Air Force Trainees	631[e]	**.00, .08,**	
	(mechanical, etc.)			**.08, .10**	t4.7
	Shipley Institute	Managers in leadership	45[f]	**.14**	t4.8
	for Living Scale	program			
	CMT–Analogies	Army platoon leaders	346[g]	.16**	t4.9
	CMT–Synonyms		344[g]	.03	t4.10
	CMT–Analogies	Army company commanders	157[h]	.17*	t4.11
	CMT–Synonyms		156[h]	.14	t4.12
	CMT–Analogies	Army battalion commanders	30[i]	.08	t4.13
	CMT–Synonyms		30[i]	.25	t4.14
Military leadership (TKML)					t4.15
Platoon leader	CMT–Analogies	Army platoon leaders	346[g]	**.18****	t4.16
Platoon leader	CMT–Synonyms		344[g]	.02	t4.17
Company commander	CMT–Analogies	Army company commanders	157[h]	**.25****	t4.18
Company commander	CMT–Synonyms		156[h]	.13	t4.19
Battalion commander	CMT–Analogies	Army battalion commanders	30[i]	**.19**	t4.20
Battalion commander	CMT–Synonyms		30[i]	.02	t4.21
Sales (TKIS)					t4.22
Global knowledge	DAT Verbal Reasoning Test	Florida State Univ. undergraduates	48[j]	.05	t4.23
Local knowledge				.40**	t4.24
Total				**n.s.**	t4.25
Global knowledge	DAT Verbal Reasoning Test	Florida State Univ. undergraduates	48[k]	−.01	t4.26
Local knowledge				.25*	t4.27
Total				**n.s.**	t4.28

n.s. = not significant. Bold entries are results where Sternberg et al. (2000) either reported the number or *correctly* reported that it was not significant. t4.29

[a] See Wagner and Sternberg (1985, p. 446). t4.30

[b] See Wagner (1987, p. 1242). Sternberg et al. (2000, p. 147) do not report the correlation, and mistakenly depict it as not significant. t4.31

[c] See Wagner and Sternberg (1985, p. 450). t4.32

[d] See Wagner (1987, p. 1244). t4.33

[e] See Sternberg et al. (2000, p. 152), who obtained the data from Eddy (1988). Although not named, the four composites are presumably similar to the Air Force's usual administrative, electronics, mechanical, and general composites. t4.34

[f] See Sternberg et al. (2000, p. 149). t4.35

[g] See Hedlund et al. (1998, p. 22). t4.36

[h] See Hedlund et al. (1998, p. 27). t4.37

[i] See Hedlund et al. (1998, p. 32). t4.38

[j] See Wagner et al. (1999, p. 163). t4.39

[k] See Wagner et al. (1999, p. 165).

* $P < .05$. t4.40

** $P < .01$. t4.41

the three Concept Mastery Tests (CMT)-*Analogy* correlations with the TKML in the Army samples, two of them significant — .18 and .25 (p. 196).[9]

Second, Sternberg et al. do not take account of restriction in range on IQ in their samples. The average for managers in leadership training on the Shipley Institute for Living Scale was IQ 120 (S.D. = 7.1), which corresponds to about the 90th percentile in the general population. Data for the two Yale undergraduate samples suggest that they are highly restricted in range on ability, and that there may also be a ceiling effect on the Differential Aptitude Test (DAT) Reasoning subtest (Form T) that all four sets of undergraduates took: specifically, the means and medians were 45–46, S.D.s were 3–4, and the range was 32–50, where 50 was the maximum possible score (Wagner, 1987, p. 1240; Wagner and Sternberg, 1985, p. 446). Restriction in range may have been similarly substantial in the three Army samples because, although there are "no known norms" for the CMT they were given, the officers' scores were comparable to ones found in an undergraduate sample (Hedlund et al., 1998, p. 22). Restriction in IQ range is probably also substantial in the other samples of workers to whom Sternberg et al. did not administer an IQ test (psychologists, managers, and sales agents), because these occupations typically recruit 70–90% of their applicants from the top half of the IQ distribution (Gottfredson, 1997, pp. 88–89). Restriction in range leads to underestimating the true correlation between tacit knowledge and IQ to some unknown extent, as Sternberg et al. (2000, p. 158) note. Like them (but for different reasons), however, I would not, in fact, expect a corrected correlation with IQ to be very high in the sorts of samples they have collected. It should be noted, however, that they actually have IQ data for only four samples of workers who took a tacit knowledge test (one of civilian managers and three of Army officers).

The more pertinent issue, however, is whether there exists a *general factor* of practical intelligence that is uncorrelated with *g*. No such evidence is ever offered. In fact, as we saw, there is no credible evidence for a general factor, let alone one uncorrelated with IQ. The best and most straightforward test of the claim that "practical intelligence is a construct that is distinct from general intelligence" would be to try to extract a general factor from a variety of tacit knowledge tests and then correlate it with IQ, or, preferably, with the *g* factor that emerges from factor analyzing broad batteries of conventional mental tests. The requisite data for such analyses do not exist.

Finally, Sternberg et al. (2000) seem to have assumed that any general factor they might discover independent of *g* would still be an *intellectual* one (another "intelligence"). The implication is that differences in knowledge must represent differences in intellectual ability or exposure to the information. That is not necessarily true, of course, because conscientiousness, interests, and other personal traits can all affect how much knowledge we seek out and accumulate on a topic. More fundamentally, however, it is not clear that Sternberg et al. have even measured knowledge as such. Recall that only on the sales test are individuals scored for their *accuracy* of response, and on all the others people are scored for the *similarity*

[9] When they get to the CMT correlations with the performance ratings, they report results sometimes for the CMT-Analogy scale and other times for the CMT-Synonym scale, but always labeling them both indistinguishably as "CMT" results (p. 197).

of their responses to those preferred by "experts." The latter procedure is more similar to the scoring of interest inventories than ability tests.

Moreover, the descriptions of the tacit knowledge tests suggest that at least some of them may capture the influence of noncognitive traits. The tacit knowledge tests for the first two samples contained a "managing career" subscale, which Wagner and Sternberg (1986, p. 56) say includes knowing "how careers are established, how reputations can be enhanced" and "how to convince others that your work is as good as it really is (or even better)." The sample items published for the tests given to the first five samples (Wagner and Sternberg, 1985, pp. 440–441, 447; Wagner, 1987, pp. 1239, 1243) do, indeed, suggest that some items on the early tacit knowledge tests focused on career advancement and tapped a calculating self-aggrandizement for impressing superiors, regardless of performance. For academic psychology, the sample items concerned the "goals... to become one of the top people in your field and to get tenure in your department." For business managers, they involved a "goal [for] rapid promotion to the top of the company" and "a chance to show your superiors what you can do in a tough situation, [with the] hope that by doing well you will improve your opportunities for advancement." Two of the eight scenarios in the more recent management test (TKIM, scenarios 6 and 8) also stress career advancement (see Sternberg et al., 2000, Appendix A). Development of the TKML explicitly excluded "self-oriented goals" when defining leadership for the study's participants (Sternberg et al., 2000, p. 177), but such goals are clearly reflected in at least half of the tacit knowledge tests.

Summary accounts by Sternberg et al. (e.g., Sternberg et al., 1995, p. 919; Sternberg et al., 2000, p. 153) of the unpublished study of managers at three levels of management (Williams and Sternberg, undated) suggest that the test used in that study may have tapped several less careerist personality traits (e.g., "how to seek out, create, and enjoy challenges" and "maintaining appropriate levels of control"). The study of 45 managers in leadership training found, however, that tacit knowledge for management seldom correlated with the scales on several personality tests, including the California Psychological Inventory (Wagner & Sternberg, 1990, p. 499).

In short, it is not clear what traits the different tacit knowledge tests may reliably tap. We certainly cannot assume that the tests' partial independence from *g* means that they measure a separate *intellectual* ability.

4.6. Empirical claim 6: practical intelligence predicts success at least as well as does academic intelligence (g)

In their preface, Sternberg et al. (2000) make it clear that their book is meant to challenge what they have described elsewhere (Sternberg & Wagner, 1993) as the "*g*-ocentric" view of intelligence. The culmination of the list of points they say the book will dispute is this:

> Moreover, practical intelligence is at least as good a predictor of future success as is the academic form of intelligence... Arguably, practical intelligence is a better predictor of success. (p. xii)

To support their claim, Sternberg et al. (2000, pp. 144–154, 196–203) summarize the findings from their six studies correlating tacit knowledge with job outcomes in five occupations: academic psychology (two samples), business management (four samples), bank management (one sample), life insurance sales (one sample), and Army officers (three samples). Sternberg et al. (2000) report 26 of the total 61 correlations of tacit knowledge scores with job outcome criteria. Shown in bold in Tables 5 and 6, the 26 range from .14 to .61, their unweighted average being .34. Sternberg et al. report another 15 correlations (with experience, age, education, etc.) in civilian samples, ranging from .26 to .41 and averaging .30 (excluding the three correlations they simply describe as "not significant").

Sternberg et al. interpret this stream of 26 criterion-related and 15 other correlations in their narrative by comparing them to the criterion validities that conventional mental tests have for predicting job performance. For example:

> These uncorrected correlations [of .2 to .4 for business managers] were in the range of the average correlation between cognitive ability test scores and job performance of .2 (Wigdor and Garner, 1982)." (Sternberg et al., 1995, p. 921)

Sternberg (1997, p. 224) translates the .2 correlation (4% of variance explained) in his book for a lay audience as "scarcely something to write home about."

The .2 estimate for *g* obviously compares unfavorably with the correlations that Sternberg et al. (2000) report for their various tacit knowledge tests. Is this contrast warranted? The answer depends on whether the comparison is accurate and appropriate. In fact, it is neither. This conclusion is based on (a) examining the size and representativeness of the samples, (b) comparing the claims for the five occupations against the data available for each, and (c) assessing the appropriateness of comparing the criterion validities for tacit knowledge tests against the suggested .2 standard for conventional tests.

5. Number, size, and representativeness of samples

The evidence that Sternberg et al. use to support their claim for the equal predictive validity of practical intelligence is meager. Although they have led readers to expect "many studies in many parts of the world with many different populations" (Sternberg et al., 2000, p. xii), a careful accounting reveals only six criterion-related studies of tacit knowledge in five occupations for a total of 11 samples of workers (Hedlund et al., 1998; Wagner, 1987; Wagner and Sternberg, 1985, 1990; Wagner et al., 1999; Williams and Sternberg, undated). As already discussed, only two of the six studies, one on civilian managers (Wagner & Sternberg, 1990) and one on Army officers (Hedlund et al., 1998), ever pit tacit knowledge against IQ in predicting job performance. And contrary to what readers have been led to believe ("we have... published most of these data ['testing our theories']... in refereed scientific journals"), only the earliest two studies are reported in peer-reviewed articles. The 1990 and 1999 publications are book chapters that only sketchily summarize unpublished work. The two remaining documents are, respectively, an unpublished 1998 technical report and an unpublished book that was cited as in press with Harcourt-Brace in 1995, with Erlbaum

Table 5
Criterion correlations for tacit knowledge in eight civilian samples

Criteria and correlates	Correlation with tacit knowledge ("total score"): academic psychologists $n = 54$[a]	$n = 91$[b]
Outcomes (in past 1[a] or 2[b] years)		
Publications (n)	**.33***	**.28*** ($n = 59$)
Citations (n)	.23	**.44***** ($n = 59$)
Conferences attended (n)	**.34***	
Conference papers presented (n)	.16	.21* ($n = 80$)
Department's scholarly rating	**.40****	**.48**** ($n = 77$)
Other		
Academic rank	− .27	
Percent time in research	**.39****	**.41***** ($n = 79$)
Percent time in teaching	− .29*	− .26* ($n = 79$)
Percent time in administration	− .41**	− .19* ($n = 79$)
Year of PhD	.04	
Age		− .22 ($n = 80$)

Criteria and correlates	Correlation with tacit knowledge ("total score"): business managers $n = 54$[c]	$n = 64$[d]
Outcomes		
Company's prestige (top of Fortune 500)	**.34***	.05 ($n = 46$)
Salary	**.46****	.21 ($n = 48$)
Level of job title	.14	
Employees supervised (n)	.10	
Other		
Management experience (years)	.21	**.30*** ($n = 49$)
Schooling beyond HS (years)	**.41****	− .01 ($n = 50$)
Age		.12 ($n = 50$)

Criteria and correlates	Correlation with tacit knowledge ("total score"): three levels of managers (n = not reported[e])
Outcomes	
Compensation	**.39*****
Age-controlled compensation	**.38*****
Level of position	**.36*****
Satisfaction	**.23***
Other	
Management experience (years)	**n.s.**
Time in position (years)	**n.s.**
Time in company (years)	− **.29****

(*continued on next page*)

Table 5 (*continued*)

Criteria and correlates	Correlation with tacit knowledge (“total score”): three levels of managers (n = not reported[e])
Companies worked for (n)	**.35*****
Higher education (years)	**.37*****
Self-reported school performance	**.26****
College quality	**.34****
Age	**n.s.**

Criteria and correlates	Correlation with tacit knowledge (“total score”): managers in leadership training (n = 45[f])
Outcome	
Average rating for two small-group managerial simulations	**.61*****

Criteria and correlates	Correlation with tacit knowledge (“total score”): bank managers (n = 29[g])
Outcomes	
Percent salary increase	**.48*** (n = 22)
Average performance rating	.37 (n = 20)
Personnel	.29 (n = 13)
New business	**.56*** (n = 13)
Policy	.39* (n = 21)

Criteria and correlates	Correlation with tacit knowledge (“total score”): life insurance salespeople (n = 48[h])		
	Total	Global	Local
Outcome			
Yearly quality awards (n)	**.35****	.25*	.28* (n=40–45)
Sales volume (1985)	.22	**.37****	−.07 (n=31)
Sales volume (1986)	.15	**.28***	−.07 (n=39)
Premiums (1985)	.20	**.26***	.02 (n=31)
Premiums (1986)	.17	**.29***	−.05 (n=39)
Other			
Time with company (years)	**.37****	.32**	.23* (n=40–45)
Times in sales (years)	**.31****	.28*	.19 (n=40–45)
Attended college	−.11	−.17	−.01 (n=40–45)
Business education	**.41****	.23	.35* (n=33)

in 2000, but for which the authors are now seeking a new publisher (Wendy Williams, personal communication, January 17, 2001).

Where Sternberg et al. (2000) offer 26 correlations to support their bold claim, personnel selection psychology offers thousands on *g,* which have in turn been extensively meta-

analyzed (e.g., Hunter, 1986; Hunter & Hunter, 1984; Schmidt & Hunter, 1998; Schmidt, Hunter, & Outerbridge, 1986; Schmidt, Hunter, Outerbridge, & Goff, 1988). Sternberg et al. provide too few samples, let alone ones with comparable outcome criteria, to perform any meta-analysis. When considering both the 26 reported and 35 unreported criterion correlations (the latter to be discussed shortly), their eight civilian samples yield the highest average criterion validities (respectively, .29, .35, .26, .13, .34, .61, .42, .18 for the eight samples in Table 5), but the samples are relatively small by personnel selection standards (average $n = 55$ for the seven with known sample sizes), meaning sampling error is high. Two of the three Army samples in Table 6 are large ($n = 163$ and 368), but all three yield very small average criterion validities (.10, .09, .10) for the relevant tacit knowledge test (TKML).

Besides the small size of most samples, they are not at all representative of people or jobs in the United States, let alone of everyday problem solving. Recall that the book's introductory claim is that "practical intelligence is at least as good a predictor of *future success* as is the academic form of intelligence" (p. xii, emphasis added). Despite the title of their book, *Practical Intelligence in Everyday Life*, Sternberg et al. report no studies of tacit knowledge for everyday tasks, not even "changing a flat tire." The patchy data on IQ's correlations with employment status, occupational level, income, crime and delinquency, welfare use, psychological adjustment, resilience, health behavior, and much more seem a cornucopia by comparison (e.g., see Brand, 1987; Gordon, 1997; Gottfredson, 1997, in press a, in press b; Herrnstein & Murray, 1994; Jencks et al., 1979; Taubman, 1977). The only nonacademic outcomes that Sternberg et al. (2000) examine relate to a very small set of fairly specific, mostly high-level occupations. They hardly represent the full range of occupations or tasks in everyday life. Moreover, all their occupations recruit individuals of above average intelligence. We clearly cannot generalize results from this tiny corner of the world to the full range of occupations and life tasks, as Sternberg et al.'s (2000) claim would have us do. For work on "common sense," it has little to do with the common man.

Notes to Table 5:

n.s. = not significant. Entries in bold are results that Sternberg et al. (2000, pp. 146–149, 151, 154, 160) report. They usually list the fuller set of variables for which data were collected in the earlier, research design sections of their narrative.

[a] See Wagner and Sternberg (1985, p. 445).

[b] See Wagner (1987, p. 1241). Scale reversed. These are "actual total" scores.

[c] See Wagner and Sternberg (1985, p. 449).

[d] See Wagner (1987, p. 1244). Scale reversed. These are "actual total" scores.

[e] See Sternberg et al. (2000, p. 154), based on Williams and Sternberg (undated). Sternberg et al. (2000, p. 154) say that the first four correlations "were computed after controlling for background and educational experience."

[f] See Wagner and Sternberg (1990, p. 498). Scale reversed. This was the only sample of civilian workers in which IQ was correlated with an outcome criterion ($r = .38$** with performance on simulated management tasks).

[g] See Wagner and Sternberg (1985, p. 451). Scale reversed.

[h] See Wagner et al. (1999, p. 166).

* $P < .05$.

** $P < .01$.

*** $P < .001$.

6. Overview of reported and unreported criterion correlations for all five occupations

I will review the criterion-related data for each occupation in turn after first noting a general problem with the reporting for all of them. The reporting of tacit knowledge's criterion-related correlations is almost always limited to the significant ones without making that fact clear. The summary narratives are such that it is very difficult for readers to remember or even know that there exist other, unreported correlations. Although the *n*-weighted average is .26 for the 22 reported correlations for which the sample size is provided (the average for the other four being .34), the *n*-weighted average for the 35 unreported correlations is .08. For the entire 57 with known sample size, the weighted average is only .15 — "scarcely something to write home about." Recall also that these are tests specifically targeted to the occupations in question.

6.1. Academic psychologists (two samples, n = 54, 91)

Sternberg et al. (2000, p. 160) summarize data for the academic psychologists as follows.

> In the field of academic psychology, correlations in the .4–.5 range were found between tacit knowledge scores and criterion measures such as citation rate, number of publications, and quality of department.

As can be seen by consulting Table 5, ".4–.5" overstates the validities of even the significant correlations they had reported earlier in the book (.28–.48; p. 146). The full set of correlations for academic psychologists ranges down to .16 and yields a weighted average of .32. Although this is clearly a respectable correlation, it is not ".4–.5."

It should also be noted that the outcome criteria for these two samples are limited to prominence in research, and relate not at all to quality of teaching or other professorial duties that would concern the employing institution. An additional problem, not mentioned by Sternberg et al. (2000), was that the response rates to the two mail surveys were very low: 18% and 28% (Wagner, 1987, p. 1239; Wagner & Sternberg, 1985, p. 441). Enhancing their appearance of scientific rigor, however, Sternberg et al. labeled these purely correlational studies as "experiments" (e.g., Sternberg & Wagner, 1993, p. 3, "more than a dozen experiments"; Wagner, 1987; Wagner & Sternberg, 1985, 1986).

6.2. Business managers (four samples; n = 54, 64, 45, and not reported, respectively)

Sternberg et al. (2000, p. 160) summarize the results as follows.

> In [two samples of] business managers, tacit knowledge scores correlated in the range of .2 to .4 with criteria such as salary, years of management experience, and whether or not the manager worked for a company at the top of the Fortune 500 list... [In a third sample, we] obtained a correlation of .61 between tacit knowledge and performance on a managerial simulation... [In a fourth sample, we] found that tacit knowledge was related to several indicators of managerial success, including compensation, age-controlled compensation, level of position, and job satisfaction, with correlations ranging from .23 to .39.

Table 6
Correlations of tacit knowledge (TKML and TKIM) and mental ability (CMT-A and CMT-S) with three levels of army officers

Nine ratings of leadership effectiveness	Platoon leaders[a] ($n=368$)				Company commanders[b] ($n=163$)				Battalion commanders[c] ($n=31$)			
	(1) TKML	(2) TKIM	(3) CMT-A	(4) CMT-S	(5) TKML	(6) TKIM	(7) CMT-A	(8) CMT-S	(9) TKML	(10) TKIM	(11) CMT-A	(12) CMT-S
Subordinate ratings												
Task	–	–	–	–	.08	**−.08**	−.12	**−.17***	.02	**.36***	**.05**	**.08**
Interpersonal	–	–	–	–	.04	**−.12**	−.16	**−.21***	−.15	**.23**	**.26**	**.31**
Overall performance	–	–	–	–	.02	**−.11**	**−.18***	**−.22****	−.02	**.24**	**.20**	**.19**
Peer ratings												
Task	.03	**.07**	**−.02**	.05	**.20***	**−.04**	−.05	−.07	–	–	–	–
Interpersonal	.03	**.00**	**−.06**	.12	.11	**−.04**	**−.20***	−.12	–	–	–	–
Overall performance	.08	**.09**	**−.05**	.04	**.19***	**−.05**	**−.18***	−.14	–	–	–	–
Superior ratings												
Task	**.14***	**−.02**	**.16***	.05	.03	**−.09**	−.04	.00	.19	**.03**	**−.04**	**−.22**
Interpersonal	**.20***	**.03**	**.09**	.03	.01	**−.15**	.01	.06	.13	**−.03**	**.27**	**.30**
Overall performance	**.14***	**−.06**	**.10**	.04	.11	**−.13**	.02	.07	**.42***	**−.07**	**.18**	**.07**
Average correlation	.10	.02	.04	.06	.09	−.09	−.10	−.09	.10	.13	.15	.12
Intercorrelations	TKML	TKIM	CMT-A	CMT-S	TKML	TKIM	CMT-A	CMT-S	TKML	TKIM	CMT-A	CMT-S
TKIM	**.36****	–	–	–	**.32****	–	–	–	**−.06**	–	–	–
CMT-Analogies	**.18****	.16*	–	–	**.25****	.17*	–	–	**.19**	.08	–	–
CMT-Synonyms	.02	.03	.41**	–	.13	.14	.61**	–	−.02	.25	.67*	–
Experience (months)	**.00**	.02	−.06	.00	**−.08**	.02	.00	−.03	**.19**	.02	−.13	−.48*

Sternberg et al. (2000) report entries in bold, although not by number if they were not significant. They only once distinguish results for CMT-A from those for CMT-S. They report variously for one or the other elsewhere in the book, but as simply "CMT."

[a] See Hedlund et al. (1998, pp. 22–23). Average $n=276$ for correlations with effectiveness; average $n=347$ for correlations among predictors.

[b] See Hedlund et al. (1998, pp. 27–28). Average $n=134$ for correlations with effectiveness; average $n=158$ for correlations among predictors.

[c] See Hedlund et al. (1998, pp. 32–33). Average $n=20$ for correlations with effectiveness; average $n=29$ for correlations among predictors.

* $P<.05$.

** $P<.01$.

The two mail surveys of business managers, again referred to as "experiments," also had low response rates: 13% and 25% (Wagner, 1987, p. 1243; Wagner and Sternberg, 1985, p. 447). There were six criterion-related correlations in all, but Sternberg et al. (2000) report only the two significant ones, which were for the first study. The second and larger of the two studies yielded no significant correlations (unless one includes irrelevant noncriteria, such as years of experience, which Sternberg et al. do in the summary quoted above). The six criterion correlations ranged from .05 to .46, their *n*-weighted average being .22. Sternberg et al.'s statement that correlations range "from .2 to .4" for these business managers, therefore, overstates the evidence, especially because the two significant correlations did not replicate for the same criteria in the parallel study.

The correlation of tacit knowledge with performance on simulated management exercises was .61 in the third sample, the study of 45 managers in leadership training (Wagner & Sternberg, 1990). This is higher than the correlation between IQ and performance (.38). Sternberg et al. (2000, p. 149) also report that the incremental validity of tacit knowledge in predicting job performance, beyond the contributions of IQ, is an additional 32% in R^2. However, these data provide at best ambiguous support for Sternberg et al.'s claim for the greater importance of tacit knowledge than IQ because the managers were already highly selected for IQ—their average was IQ 120. When most differences in IQ have been eliminated in a sample, the small remaining differences have little power to predict anything. It may, therefore, falsely appear that IQ is much less powerful than other (nonrestricted) variables, even when it is much more powerful in representative samples. In any case, this study has never been described in much detail, making evaluation difficult.

The correlations that Sternberg et al. (2000, p. 154) report for the fourth study, of managers at three levels (Williams and Sternberg, undated), range from .23 for satisfaction to .39 for compensation. As additional support, they report that tacit knowledge increased R^2 by .04 and .05, respectively, for "maximum compensation" and "maximum compensation controlling for age" after controlling for some combination of age, education, and experience. It is impossible to evaluate the study, however, or to know what the unreported correlations are, because there is no further information available on this long-in-press study, even the sample size. Moreover, none of its reported criteria relate to actual job performance.

6.3. Bank managers (one sample; n = 29)

Sternberg et al. (2000, p. 160) do not describe this study in their book, but they do summarize its results with a sentence in their conclusion on civilian studies:

> In a study with bank branch managers, [we] obtained significant correlations between tacit knowledge scores and average percentage of merit-based salary increase ($r = .48$, $P < .05$) and average performance rating for the category of generating new business for the bank ($r = .56$, $P < .05$).

The correlations they mention are for two of the three significant ones out of a total five. Although the correlations they mention are high (.46 and .58), the weighted average for all

five is somewhat lower, .42. The study was extremely small, however, with *n*s ranging from 13 to 22 for the individual criteria (Wagner & Sternberg, 1985, p. 451).

6.4. Life insurance sales (one sample, n = 48)

The summary sentence is as follows.

> In studies with salespeople, Wagner, Rashotte, & Sternberg (1994) found correlations in the .3 to .4 range between tacit knowledge and criteria such as sales volume and sales awards received. (Sternberg et al., 2000, p. 160)

A look at the prior, more extensive published account of this study (Wagner et al., 1999) shows that the criterion correlations varied considerably depending on whether the tacit knowledge was "local" (− .07–.28), "global" (.25–.37), or combined into a "total" score (.15–.35). Sternberg et al. (2000, p. 160) report criterion correlations for the total score only when they were significant (one was), and, when not, report them for global knowledge (all of whose five correlations were significant whereas only one of the unreported "local" ones was). The claim of ".3–.4" overstates the results somewhat even for the eight significant ones, because those correlations ranged from .25 to .37 before rounding. The average for all 15 was considerably lower: .18.

6.5. Army leadership (three samples; n = 368, 163, 31, respectively)

Hedlund et al. (1998) correlated tacit knowledge scores for the three levels of leadership with six to nine performance ratings for each officer (task, interpersonal, and overall by subordinates, peers, and superiors). What follows is virtually the entire published account of the criterion-related results of that 6-year study to understand tacit knowledge's role in military leadership:

> At all three levels, we obtained evidence of convergent validity of the TKML with LES [Leadership Effectiveness Survey] ratings. The pattern of these relationships varied across rater sources and across levels. At the platoon level, higher TKML scores correlated significantly with higher effectiveness ratings by superiors on all three leadership dimensions (*r*'s of .14 to .20, $P < .05$). At the company level, higher TKML scores correlated significantly with higher effectiveness ratings by peers for overall and task leadership (*r*'s of .19 and .20 respectively, $P < .05$). At the battalion level, higher TKML scores correlated significantly with higher ratings of overall effectiveness by superiors ($r = .42$, $P < .05$). (Sternberg et al., 2000, p. 198)

As is true for the book's summaries of all the other criterion studies, it is very difficult to discern what the full set of criterion correlations is that Sternberg et al. (2000) are drawing from. Table 6, therefore, reproduces it from the pertinent Army technical report (Hedlund et al., 1998; excluding the criterion correlations for experience). In their summary of these unpublished results, Sternberg et al. (p. 198) mention six correlations ranging from .14 to .42, the unweighted average being .22. These are, however, only the six significant correlations out of the full 21 for the tacit knowledge test targeted to the three jobs in question (the three versions of the TKML). As seen in Table 6, the correlations for the three samples average .10,

.09, and .10 (columns 1, 5, and 9 in the top panel), for a weighted average of .10. Recall that this is by far the largest criterion-related study of tacit knowledge.

Turning to the correlations of IQ (CMT-A and CMT-S) with performance ratings (columns 3–4, 7–8, and 11–12 in the top panel of Table 6), Sternberg et al. (2000, p. 197) report all seven significant correlations, pointing out that six of the seven are negative. The averages of the CMT correlations with performance for the three samples are .05, −.09, and .14, for a weighted average of −.01. On the surface, this comparison of average criterion validities (.10 for the TKML and −.01 for the CMT) would seem to favor the TKML. One is given pause, however, by the fact that all the negative correlations for the CMT were from a single sample, company commanders (columns 5–8), and that sample's results were peculiar. As can be seen in Table 6, performance correlated almost uniformly negatively with all predictors except the TKML (column 5) in that sample—the TKIM, the two CMT tests, and experience (not shown here; see Hedlund et al., 1998, p. 28, for data on experience). In any case, there was "nothing to write home about" for either tacit knowledge or IQ in this large-scale study.

The Army study is not only the largest tacit knowledge study, but also the only one to have administered two different tacit knowledge tests to the same sample of workers in addition to measuring IQ. It, therefore, provides the best single test of Sternberg et al.'s (2000) claim that a *general factor* of practical intelligence predicts performance as well as does IQ. They do not use the study for this purpose, however, but focus instead on whether the more relevant test, the TKML, adds predictive value above and beyond that afforded by the less relevant tacit knowledge test (the TKIM) as well as IQ (the two CMT tests). Answering this question, Sternberg et al. (2000, p. 199) report as additional support for the TKML that it significantly increased the amount of variance explained in 2 of the 5 pertinent sets of ratings (peer and superior ratings for platoon leaders and peer, superior, and subordinate ratings for company commanders): namely, increases in R^2 ranged between .02 and .04 for task, interpersonal, and overall performance for platoon leaders (for ratings by superiors only), and .03–.06 for the three performances for company commanders (for ratings by peers only).

Whatever thin support this might provide for the "domain-*specificity*" of the TKML, it provides none at all for its "domain-*generality*," that is, for the validity of any *common factor* that it shares with the TKIM. As the unpublished technical report (Hedlund et al., 1998, pp. 24, 29) reveals, the TKIM never accounted for a significant amount of the variance in performance ratings, meaning that whatever it shared in common with the TKML also failed to predict performance. This failure can be seen in the simple correlations in six for the two samples in question (platoon leaders and company commanders). In none of the 15 opportunities (columns 2 and six in Table 6) did the TKIM correlate significantly with job performance. Turning to battalion commanders, in the one case where the TKIM did predict performance (column 10), the TKML did not (column 9). In other words, never did both tests significantly predict the same performance rating. If there is a common factor, it was too weak to predict performance in this fairly large study.

It is not clear why the study's correlations tended to be so surprisingly low. It is clear, however, that they provide no support whatever for a general factor of practical intelligence that rivals *g* in practical importance.

7. Aptness of the contest between tacit knowledge and IQ, including the .2 average criterion validity for conventional tests

Although Sternberg et al. treat all the studies as equally pertinent to testing their claims, there are reasons beyond sample size to accord some studies greater consideration than others. One concerns the criteria used to validate the tacit knowledge tests. They were of two very different types: career advancement and on-the-job performance. Practical intelligence theory does not clearly distinguish the two, sometimes stressing one and sometimes the other. Sternberg et al. simply lump the two sorts of outcome measures together (also with *predictors*, such as education and experience) as "criteria" or "criterion reference measures" (e.g., Wagner, 1987, p. 1239). When the concern is job performance, as it is in Sternberg et al.'s (2000) test of empirical claim 6, the careerist outcomes are not relevant. When the concern is life success, such as income, they are. However, that would require comparing the validities for tacit knowledge with sociological data relating IQ to income, which Sternberg et al. fail to do.

The first 5 samples listed in Table 5 used primarily careerist criteria (salary, level of title, eminence of department, working in a top Fortune 500 company, satisfaction, and the like). They, therefore, do not seem relevant in testing empirical claim 6. The remaining three samples in Table 5 and the three Army samples in Table 6 are more relevant to testing the claim, because they used mostly job performance criteria (performance ratings, sales volume, and the like), although the procedures for getting ratings are clear only for the Army samples (Sternberg et al., 2000, pp. 189–190, 192). *n*-weighted criterion correlations for the five samples with career-oriented criteria averaged .28 (excluding the sample with unreported sample size, whose average for *reported* correlations was .34) vs. .12 for the six samples with performance-oriented criteria (.24 with the Army data excluded). The criterion validities for the more relevant criteria are, thus, half those of the less relevant (.12 vs. .28).

A second problem with Sternberg et al.'s (2000) comparison of tacit knowledge with *g* is that the .2 average correlation they accord IQ is false. Nowhere does their cited source, the National Academy of Science (NAS; Wigdor & Garner, 1982), say that the average correlation between cognitive ability test scores and job performance is .2. The proffered .2 average (Sternberg et al., 1995, p. 921) seems to refer to a number that the NAS committee specifically rejected in favor of a higher average correlation. This is what the committee (Wigdor & Garner, 1982, p. 142) actually said when reviewing research on the criterion validities for cognitive tests:[10]

> Ghiselli summarized his work as indicating that for all occupations, the average validity of employment tests for... proficiency on the job [is] .19... It is probable that Ghiselli's average figures are somewhat lower than the coefficients a survey of current test use would provide... Ghiselli himself did a second, smaller study of standardized tests used in

[10] The only other possible source for the .2 number occurs when the NAS committee explains what a correlation coefficient is. One of the examples it provides is that "correlations of only about .2 are fairly common for occupational performance measures" (Wigdor & Garner, 1982, p. 56).

> personnel selection in 1973; for 21 job categories, he reported average validities of. . . .35 for job performance criteria.

A later report by the NAS on the US Department of Labor's General Aptitude Test Battery (GATB; Hartigan & Wigdor, 1989, p. 5) is consistent on this point with the earlier one: "In the 750 studies, the correlations of GATB-based predictors with supervisor ratings, after correction for sampling error, are in the range of .2 to .4."

Even this .2–.4 range probably underestimates the average uncorrected validity for jobs in general, because the GATB was used to screen disproportionately for lower-level industrial jobs. Large-scale validation research on the Armed Forces Qualification Test (AFQT) has routinely found uncorrected criterion correlations of .3–.6 with job performance in mid-level jobs (Sticht, 1975; Wigdor & Green, 1991). The huge Joint-Service Job Performance Measurement Project (JPM)—again, reviewed by the NAS—found that the median correlation of hands-on (i.e., objective job sample) performance with the AFQT was .38 for the 23 high volume jobs studied, with the AFQT predicting later performance equally well in all four military services (Wigdor & Green, 1991, p. 161). (The JPM study, unlike Sternberg et al.'s various studies, measured IQ prior to job entry.) Uncorrected AFQT predictions of hands-on performance in the four Marine jobs studied reinforce the point that the supposedly academic AFQT predicts performance in nonacademic jobs surprisingly well: rifleman (.55), machine gunner (.66), mortarman (.38), and assaultman (.46; Wigdor & Green, 1991, p. 161).

The average criterion validity among civilian jobs in the United States, after correcting for unreliability of measurement and restriction in range, is about .5 (Schmidt & Hunter, 1998). Although it is appropriate for Sternberg et al. (2000) to argue that the uncorrected correlations for tacit knowledge ought to be compared to uncorrected correlations for IQ, Sternberg's (1997, p. 225, emphasis in original) glib aspersions on these routine statistical corrections are not:

> Some psychologists. . . have suggested that the validity coefficient of IQ tests and related measures for predicting job performance is really about .5, not .2. That's a pretty big difference. How did they get a figure so much higher than that reported by the commission appointed by the prestigious National Academy of Science? They used a variety of what euphemistically might be called *statistical corrections* in order to jack up these validity coefficients.

Actually, professional test standards (Society for Industrial and Organizational Psychology, 1987, Standard B.5.b, p. 16) and good test practice (Cronbach, 1990, pp. 213–214, 432–433) *require* that correlations be corrected for some purposes, the required corrections differing by specific purpose. The greatest number of statistical corrections is required for the present purpose, namely, theory testing (e.g., the relation between underlying constructs). It is Sternberg et al. (2000) who have surreptitiously "jacked up" their results, because they consistently cite only the highest correlations for tacit knowledge while understating those for IQ. Selectively reporting only the highest correlations is not a way to correct for bias; rather, it creates bias.

In any case, it is not appropriate for Sternberg et al. to compare the correlations for tacit knowledge in mostly mid- to high-level jobs with those for *g* in all jobs. Recall that the predictive validity of *g* rises with job complexity level. If Sternberg et al. wish to compare the correlations for tacit knowledge with analogous correlations for IQ, the appropriate compar-

isons would be with like occupations, specifically, other moderate- to high-level jobs in management and leadership. Whatever the outcome, it could not be generalized too broadly, however, because the socioemotional/motivational dimensions of job performance depend more on noncognitive, personality traits than do the more strictly instrumental dimensions of work (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990), and these jobs seem to stress the former sorts of duties more than do most jobs.

Finally, even if the correlations for tacit knowledge were as high as or higher than those for *g* in comparable occupations, that would still provide no evidence for an equally important "general factor of practical intelligence." Conventional mental tests are largely interchangeable for purposes of measuring *g* and, thus, it matters little which particular one is used to predict job performance, as long as it is reliable and highly *g* loaded. They all measure the same active ingredient—*g*. In contrast, Sternberg et al.'s tacit knowledge tests are specific to particular jobs: "tacit knowledge is always wedded to particular uses in particular situations or in classes of situations" (Sternberg et al., 1995, p. 917). Each job, even each job level (as in their Army and three-levels-of management studies), therefore, needs its own targeted tacit knowledge test. Moreover, if the aim is to select better workers, such tests can probably be fairly administered only to people who are already experienced. This is the case with all job knowledge tests. If there really is a practical intelligence (a general factor for "common sense") that is comparable to the general factor for "academic" intelligence (*g*), then Sternberg et al. should be able to create a test or extract a common factor from a set of them that has predictive validity in many *different* settings, as does *g*. They have not done so, nor have they said they will.

The contest they have set up is a false one. It is akin to saying that I can keep up with you in any sport, but then I bring in my brother to run the track meets, my sister to compete in tennis, my dad in golf, and my cousin in swimming, while you must compete in all of them yourself. Where I may offer different specific forms of highly cultivated expertise, you must possess an all around ability to compete in any sport, practiced or not. However, if I really wish to support my claim that I possess a different but equally powerful general ability than you do, I must compete in all those sports myself. This requirement for our contest does not imply that practice, experience, and expertise are unimportant. Far from it. It just means that no form of "developing expertise" is comparable to a general ability, such as *g,* at either a conceptual or empirical level. Precisely because tacit knowledge is expertise, it is specific and not general, and Sternberg et al. (2000) have provided no evidence for a general factor of tacit knowledge that transcends this specificity and, thus, represents a practical intelligence with broad predictive value. Conversely, labeling IQ as only one form of developing expertise, as Sternberg et al. do, does not erase the general factor of intelligence, *g*, and its broad predictive value in jobs and beyond.

8. Conclusions

Sternberg et al. have made an implausible claim, namely, that tacit knowledge reflects a general factor of intelligence that equals or exceeds *g* in its generality and everyday utility.

They back it up mostly with the appearance, not the reality, of hard evidence. The foregoing examination of their evidence has shown how they appear to play the scientific game more than they really do; that the "reputation they build is not tantamount to the quality of the work."

The authors of *Practical Intelligence in Everyday Life* first ask us to suspend belief on the evidence that is plain to see for all who would look: in particular, the massive evidence from many decades of research that reveals *g* to be a highly general mental ability with strong genetic roots that distinguishes among us in socially important ways. Their book then asks us to accept its meager data as firm evidence for a coequal, if not *more* general and *more* useful, practical intelligence: in particular, their odd collection of examples and anecdotes of mostly ill-educated people succeeding at mostly simple tasks they have practiced extensively, and their small number of usually small samples of brighter-than-average workers whose differences in "knowing the ropes" in their mostly high-level jobs help predict how well they perform their jobs or get ahead in them.

Their various summary reports (e.g., Sternberg et al., 1995, 2000), which contain the only published information for several of the six studies, also exaggerate the strength of the empirical support they summarize. They do so by selectively presenting the most favorable results; overstating even those; interpreting inconsistent data in ways that produce consistent support; and giving citations to back up strong statements but which do not actually provide independent support (many are just earlier summaries of the same thing) or that even contradict the claim in question.

The authors simultaneously discourage the close analysis that would reveal the inadequacies of their data and presentation. They do so partly by appealing to many people's strong desire to believe them, specifically, by tapping the popular preference for an egalitarian plurality of intelligences (everyone can be smart in some way) and a distaste for being assessed, labeled, and sorted by inscrutable mental tests. These sentiments are evoked again by casting aspersions on research and researchers that have helped reinstate the concept of *g*, or *general* intelligence.

It is true that *g* provides only a partial explanation of "intelligent behavior," and that its role in everyday affairs is yet poorly understood. But there is a solid, century–long evidentiary base upon which researchers are busily building. Simply positing a new and independent intelligence to explain much of what remains unexplained (and much of what has *already* been explained), while simultaneously ignoring the ever-growing evidentiary base, does not promise to advance knowledge. The concept of tacit knowledge does, I suspect, point to a form of experience and knowledge that lends itself to the development of what might be called wisdom — a gradual understanding of the probabilities and possibilities in human behavior (and in individual persons) that we generally develop only by experiencing or observing them first-hand over the course of our lives. This is not a new form of intelligence, however, but perhaps only the motivated and sensitive application of whatever level of *g* we individually possess. Sternberg et al. could better advance scientific knowledge on this issue by probing more deeply and analytically into the role of tacit knowledge in our lives rather than continuing to spin gauzy illusions of a wholly new intelligence that defies the laws of evidence.

9. Uncited references

Sternberg & Wagner, 1986

Acknowledgments

I wish to thank reviewers Earl B. Hunt, Arthur R. Jensen, and Philip A. Vernon, as well as Jan H. Blits, Robert A. Gordon, and David Lubinski, for their very helpful comments on earlier versions of this paper.

References

Arvey, R. D. (1986). General ability in employment: a discussion. *Journal of Vocational Behavior*, *29* (3), 415–420.

Baldwin, J., Kirsch, I. S., Rock, D., & Yamamoto, K. (1995). *The literacy proficiencies of GED examinees: results from the GED–NALS comparison study*. Washington, DC: American Council on Education and Educational Testing.

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

Brand, C. (1987). The importance of general intelligence. In: S. Modgil, & C. Modgil (Eds.), *Arthur Jensen: consensus and controversy* (pp. 251–265). New York: Falmer.

Brand, C. (1996). *The g factor: general intelligence and its implications*. Chichester, England: Wiley.

Brody, N. (in press). Construct validation of the Sternberg Triarchic Abilities Test (STAT): comment and reanalysis. *Intelligence*.

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

Carroll, J. B. (1993). *Human cognitive abilities: a survey of factor-analytic studies*. New York: Cambridge University Press.

Casto, S. D., DeFries, J. C., & Fulker, D. W. (1995). Multivariate genetic analysis of Wechsler Intelligence Scale for Children-Revised (WISC-R) factors. *Behavior Genetics*, *25* (1), 25–32.

Cattell, R. B. (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 (Ed.), *Practical intelligence: nature and origins of competence in the everyday world*. New York: Cambridge University Press.

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

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist*, *30*, 15–41.

Colombo, J. (1993). *Infant cognition: predicting later intellectual functioning*. Newbury Park, CA: Sage.

Cornelius, S. W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging*, *2*, 144–153.

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

Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional intelligence: in search of an elusive construct. *Journal of Personality and Social Psychology*, *75*, 989–1015.

Deary, I. J. (2000). *Looking down on human intelligence: from psychometrics to the brain*. Oxford: Oxford University Press.

Denney, N. W., & Palmer, A. M. (1981). Adult age differences on traditional and practical problem-solving measures. *Journal of Gerontology*, *36*, 323–328.
Dörner, D., & Kreuzig, H. (1983). Problemlösefähigkeit und Intelligenz. *Psychologische Rundschaus*, *34*, 185–192.
Dörner, D., Kreuzig, H., Reither, F., & Staudel, T. (1983). *Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.
Eddy, A. S. (1988). *The relationship between the Tacit Knowledge Inventory for Managers and the Armed Services Vocational Aptitude Battery.* Unpublished master's thesis. St. Mary's University, San Antonio, TX.
Epstein, R. (1999, November/December). You're smarter than you think. *Psychology Today*, p. 30.
Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences*. New York: Basic.
Goleman, D. (1995). *Emotional intelligence*. New York: Bantam.
Gordon, R. A. (1997). Everyday life as an intelligence test: effects of intelligence and intelligence context. *Intelligence*, *24* (1), 203–320.
Gottfredson, L. S. (1997). Why *g* matters: the complexity of everyday life. *Intelligence*, *24* (1), 79–132.
Gottfredson, L. S. (in press a). *g*: highly general and highly practical. In R. J. Sternberg, E. L. Grigorenko (Eds.), *The general intelligence factor: how general is it?* Mahwah, NJ: Erlbaum.
Gottfredson, L. S. (in press b). *g*, jobs, and life. In H. Nyborg (Ed.), *The scientific study of intelligence: tribute to Arthur R. Jensen.* Permagon.
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.
Hartigan, J. A., Wigdor, A. K. (Eds.) (1989). *Fairness in employment testing: validity generalization, minority issues, and the general aptitude test battery.* Washington, DC: National Academy Press.
Hedlund, J., Horvath, J. A., Forsythe, G. B., Snook, S., Williams, W. M., Bullis, R. C., Dennis, M., Sternberg, R. J. (1998, April). *Tacit knowledge in military leadership: evidence of construct validity.* Technical Report 1080. Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.
Herrnstein, R. J., & Murray, C. (1994). *The bell curve: intelligence and class structure in American life*. New York: Free Press.
Hogan, R. T. (1991). Personality and personality measurement. In: M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., vol. 2, pp. 873–919). Palo Alto, CA: Consulting Psychologists Press.
Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallised general intelligences. *Journal of Educational Psychology*, *57*, 253–270.
Humphreys, L. G. (1986). Commentary. *Journal of Vocational Behavior*, *29* (3), 421–437.
Hunt, E. (1995). *Will we be smart enough? A cognitive analysis of the coming workforce*. New York: Russell Sage Foundation.
Hunt, E. (2001). Multiple views of multiple intelligence *[review of intelligence reframed: multiple intelligence in the 21st century]*. *Contemporary Psychology*, *46* (1), 5–7.
Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29* (3), 340–362.
Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*, 72–98.
Jencks, C., Bartlett, S., Corcoran, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., Mueser, P., Olneck, M., Schwartz, J., Ward, S., & Williams, J. (1979). *Who gets ahead? The determinants of economic success in America*. New York: Basic Books.
Jensen, A. R. (1993). Test validity: *g* versus "tacit knowledge". *Current Directions in Psychological Science*, *1*, 9–10.
Jensen, A. R. (1998). *The g factor: the science of mental ability.* Westport, CN: Praeger.
Kirsch, I. S., Jungeblut, A., Jenkins, L., & Kolstad, A. (1993). *Adult literacy in America: a first look at the results of the National Adult Literacy Survey.* Princeton, NJ: Educational Testing Service.
Kline, P. (1991). Sternberg's components: non-contingent concepts. *Personality and Individual Differences*, *12* (9), 873–876.

Kline, P. (1998). *The new psychometrics: science, psychology and measurement*. London: Routledge.
Lave, J., Murtaugh, M., & de la Roche, O. (1984). The dialectic of arithmetic in grocery shopping. In: B. Rogoff, & J. Lace (Eds.), *Everyday cognition: its development in social context* (pp. 67–94). Cambridge, MA: Harvard University Press.
Lichtenstein, P., & Pedersen, N. L. (1997). Does genetic variance for cognitive abilities account for genetic variance in educational achievement and occupational status? A study of twins reared apart and twins reared together. *Social Biology*, *44* (1–2), 77–90.
Lubinski, D., & Benbow, C. P. (1995). An opportunity for empiricism: review of Howard Gardner's Multiple intelligences: the theory in practice. *Contemporary Psychology*, *40*, 935–938.
Lubinski, D., & Dawis, R. V. (1992). Aptitudes, skills, and proficiencies. In: M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (vol. 3, pp. 1–59). Palo Alto, CA: Consulting Psychologists Press.
Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, *24* (1), 159–201.
McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. S., & Ashworth, S. (1990). The relationship between predictor and criterion domains. *Personnel Psychology*, *43* (2), 335–366.
Messick, S. (1992). Multiple intelligences or multilevel intelligence? Selective emphasis on distinctive properties of hierarchy: on Gardner's Frames of Mind and Sternberg's Beyond IQ in the context of theory and research on the structure of human abilities. *Psychological Inquiry*, *3*, 365–384.
Moffitt, T. E., Caspi, A., Harkness, A. R., & Silva, P. A. (1993). The natural history of change in intellectual performance: who changes? How much? Is it meaningful? *Journal of Child Psychology and Psychiatry*, *34*, 455–506.
Murtaugh, M. (1985). The practice of arithmetic by American grocery shoppers. *Anthropology and Education Quarterly*, *16*, 186–192.
Plomin, R., & Bergman, C. S. (1991). Nature and nurture: genetic influences on "environmental" measures. *Behavioral and Brain Sciences*, *14*, 373–427.
Plomin, R., & DeFries, J. C. (1998). The genetics of cognitive abilities and disabilities. *Scientific American*, *278* (5), 62–67.
Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2000). *Behavioral genetics* (4th ed.). New York: Worth Publishers and W.H. Freeman.
Rabbitt, P. (1988). Human intelligence: a critical review of five books by R.J. Sternberg. *The Quarterly Journal of Experimental Psychology*, *40A* (1), 167–185.
Ree, M. J., & Earles, J. A. (1993). *g* is to psychology what carbon is to chemistry: a reply to Sternberg and Wagner, McClelland, and Calfee. *Current Directions in Psychological Science*, *2* (1), 11–12.
Schmidt, F. L., & Hunter, J. E. (1993). Tacit knowledge, practical intelligence, general mental ability, and job knowledge. *Current Directions in Psychological Science*, *2* (1), 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* (2), 262–274.
Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, *71* (3), 432–439.
Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relations of experience and ability with job performance: test of three hypotheses. *Journal of Applied Psychology*, *73* (1), 46–57.
Science and pseudoscience (1999, July/August). *APS Observer*, p. 27.
Scribner, S. (1984). Studying working intelligence. In: B. Rogoff, & J. Lave (Eds.), *Everyday cognition: its development in social context* (pp. 9–40). Cambridge, MA: Harvard University Press.
Scribner, S. (1986). Thinking in action: some characteristics of practical thought. In: R. J. Sternberg, & R. K. Wagner (Eds.), *Practical intelligence: nature and origins of competence in the everyday world* (pp. 13–30). New York: Cambridge University Press.
Society for Industrial and Organizational Psychology (1987). *Principles for the validation and use of personnel selection procedures* (3rd ed.). College Park, MD: Author.

Sternberg, R. J. (1985). *Beyond IQ: a triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In: M. G. KcKeown, & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89–105). Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1988). *The triarchic mind: a new theory of human intelligence*. New York: Viking.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (2000). Human intelligence: a case study of how more and more research can lead us to know less and less about a psychological phenomenon, until finally we know much less than we did before we started doing research. In: E. Tulving (Ed.), *Memory, consciousness, and the brain: the Tallinn Conference* (pp. 363–373). Philadelphia, PA: Taylor and Francis, Psychology Group.

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conception of intelligence. *Journal of Personality and Social Psychology*, *41*, 37–55.

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.

Sternberg, R. J., & Kaufman, J. C. (1998). Human abilities. *Annual Review of Psychology*, *49*, 479–502.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, *38*, 878–893.

Sternberg, R. J., Wagner, R. K. (Eds.) (1986). *Practical intelligence: nature and origins of competence in the everyday world*. New York: Cambridge University Press.

Sternberg, R. J., & Wagner, R. K. (1993). The *g*-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science*, *2* (1), 1–5.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: the nature and role of tacit knowledge in work and at school. In: J. M. Puckett, & H. W. Reese (Eds.), *Mechanisms of everyday cognition* (pp. 205–227). Hillsdale, NJ: Erlbaum (Note: this is the article that Sternberg et al., 2000, cite as being published in H. Reese and J. Puckett, *Advances in lifespan development*).

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist*, *50* (11), 912–927.

Sticht, T. (Ed.) (1975). *Reading for working: a functional literacy anthology*. Alexandria, VA: Human Resources Research Organization.

Tambs, K., Sundet, J. M., Magnus, P., & Berg, K. (1989). Genetic and environmental contributions to the covariance between occupational status, educational attainment, and IQ: a study of twins. *Behavior Genetics*, *19 (2), 209–222.*

Taubman, P. (Ed.) (1977). *Kinometrics: determinants of socioeconomic success within and between families*. New York: North-Holland Publishing.

Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: genetic overlap but environmental differences. *Psychological Science*, *2* (3), 158–165.

Vernon, P. E. (1971). *The structure of human abilities* (3rd ed.). New York: Wiley.

Vineberg, R., Taylor, E. N. (1972). Performance in four Army jobs by men of different aptitude (AFQT) levels: 3. The relationship of AFQT and job experience to job performance. *Human Resources Research Organization Technical Report 72-22.* Washington, DC: Chief of Research and Development, US Department of the Army.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology*, *52* (6), 1236–1247.

Wagner, R. K., Rashotte, C. A., Sternberg, R. J. (1994). Tacit knowledge in sales: rules of thumb for selling to anyone. *Paper presented at the annual meeting of the American Educational Research Association,* Washington, DC.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: the role of tacit knowledge. *Journal of Personality and Social Psychology*, *48*, 436–538.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In: R. J. Sternberg, R. K. Wagner (Ed.), *Practical intelligence: nature and origins of competence in the everyday world* (pp. 51–83). New York: Cambridge University Press.

Wagner, R. K., & Sternberg, R. J. (1990). Street smarts. In: K. E. Clark, & M. B. Clark (Eds.), *Measures of leadership* (pp. 493–504). Clark Orange, NJ: Leadership Library of America.

Wagner, R. K., Sujan, H., Sujan, M., Rashotte, C. A., & Sternberg, R. J. (1999). Tacit knowledge in sales. In: R. J. Horvath, & J. A. Horvath (Eds.), *Tacit knowledge in professional practice: researcher and practitioner perspectives* (pp. 155–182). Mahwah, NJ: Erlbaum.

Wigdor, A., Garner, W. R. (Eds.) (1982). *Ability testing: uses, consequences, and controversies: Part I. Report of the committee*. Washington, DC: National Academy Press.

Wigdor, A. K., Green Jr., B. F. (Eds.) (1991). *Performance for the workplace* (vol. 1). Washington, DC: National Academy Press.

Williams, M. V., Baker, D. W., Parker, R. M., & Nurss, J. R. (1998). Relationship of functional literacy to patients' knowledge of their chronic disease. *Archives of Internal Medicine*, *158*, 166–172.

Williams, S. A., Denney, N. W., & Shadler, M. (1983). Elderly adults' perception of their own cognitive development during the adult years. *International Journal of Aging and Human Development*, *16*, 147–158 (Get and check whether same sample as earlier study).

Williams, W. M., Sternberg, R. J. (undated). *Success acts for managers.* Ithaca, NY: Cornell University (note: Sternberg et al., 1995, listed the book as in press with Harcourt-Brace; Sternberg et al., 2000, as in press with Erlbaum Publishers; and the authors are now looking for a new publisher [Wendy Williams, personal communication, January 17, 2001]).

Willis, S. L., & Schaie, K. W. (1986). Practical intelligence in later adulthood. In: R. J. Sternberg, & R. K. Wagner (Eds.), *Practical intelligence: nature and origins of competence in the everyday world* (pp. 236–268). New York: Cambridge University Press.