Cognitively SPEAKING

An update on Ability Measurement with *CogAT*®
Volume #4 • SUMMER Issue • 2005

Riverside Publishing
A HOUGHTON MIFFLIN COMPANY

IS ABILITY SEPARATE FROM ACHIEVEMENT?

by Dr. David Lohman, The University of Iowa

I sat there squirming as he spoke. The presenter, the author of a nonverbal reasoning test, proceeded to ridicule today's most respected ability tests because some items on them require knowledge of language or other learned symbol systems. The *Cognitive Abilities Test™ (CogAT®)* was one of the tests singled out for criticism. "How could one possibly estimate the ability of a child by asking her to read and complete sentences or to make judgments about the relative size of 2 + 3 and 2 x 3? Aren't these learned skills?" A good ability test, the presenter claimed, should be one that could be administered with equal fairness to all individuals. The abilities required to solve items should be independent of culture (especially language) and experience (especially education).

Is it possible to separate innate ability from the knowledge and skills acquired through language, culture, and schooling? In this issue of *Cognitively Speaking*, I try to show that (1) the belief that ability is innate and achievement is acquired is not based on scientific theory; (2) the development of beliefs about ability and achievement follows a predictable sequence; and (3) better understanding of these constructs can improve both the process by which gifted students are identified and the programs designed to serve them.

Confusion over the relationship between ability and achievement is not new. In 1927 Truman Kelley, the lead author of the *Stanford Achievement Test,* expressed alarm when people treated his achievement test and intelligence tests as if they measured independent constructs. Kelley knew that the overlap on ability and achievement tests was enormous. The culprit, he said, was language. Different names are used for intelligence tests and achievement tests, so we expect that they measure different things.

A Theory about Ability and Achievement

Almost everyone starts out believing that ability and achievement are separate. Then, as their expertise grows, some develop more sophisticated theories. Why some move on while others retain their original beliefs is not always clear. One factor that seems to matter is a willingness to consider evidence that contradicts one's current views. Openness to a new perspective is difficult if one has a vested interest in preserving a current belief system. Cognitive styles may also matter. A fondness for the sort of distinct categories that typify basic theories may make it difficult to move on to worlds in which there is more gray than black and white. The progression goes something like this:

> Some individuals encounter evidence that challenges their beliefs in the independence of intelligence tests and achievement tests.

1. Things Are What They Seem to Be.

Virtually everyone starts here. The individual believes that ability tests measure (or ought to measure) innate potential. This means that culture, education, personal experience, and motivation should not influence scores. Similarly, achievement tests measure (or ought to measure) only knowledge and skills learned in school. Individuals retain these beliefs unless they encounter new evidence.

CogAT®

2. New Evidence Appears.

Some individuals encounter evidence that challenges their beliefs in the independence of intelligence tests and achievement tests. As in Kelley's (1927) case, this evidence may come from statistics that show high correlations between the two measures. It may also come from an inspection of ability and achievement tests that shows similarities in content and structure.

Reactions to this knowledge take several forms. Some conclude that the overlap reflects mostly ability (Spearman, 1923; Jensen, 1998). Others look at the same data and say that the overlap is mostly the product of learning (Ferguson, 1956; Humphreys, 1981; Thorndike, et al. 1926). Neither of these responses has much to say about the differences between ability and achievement.

A popular solution that attends to both the similarity and difference between ability and achievement tests is to create a continuum on which tasks vary by their novelty. The more achievement-like tasks are placed at the low-novelty end of the continuum. The more ability-like tasks are placed at the high-novelty end. A continuum like this may be found in the writings of Stern (1914), Thorndike et al. (1926), Anastasi (1937), Cattell (1943; 1963), Cronbach (1970), Snow (1980), and Sternberg (1985), to name a few. Placing both types of tasks on the same continuum recognizes their commonality. Placing them at opposite ends also recognizes their uniqueness.

Interpretations of this continuum vary. For example, in his original proposal of the theory of fluid and crystallized abilities, Cattell (1943) emphasized the equality of these two intelligences. In later versions of the theory, however, Cattell (1963, 1971) interpreted fluid ability as something like the individual's true, innate intelligence that when invested in experience produced a particular constellation of crystallized abilities. However, there is no evidence that fluid abilities are more influenced by biological factors than crystallized abilities. Further, contrary to theory, crystallized abilities generally have a greater impact on the development of fluid abilities than vice versa (Horn & Noll, 1997).

The attempt to interpret fluid reasoning abilities as the real intelligence is perhaps better understood as an attempt to preserve one's original beliefs in the face of new evidence. There are other examples of this tendency. An extreme version of this view is the claim that nonverbal tests are a pure measure of abilities. According to this view, tests that measure reasoning abilities should not be contaminated by content or skills that would influence performance on an achievement test. Conversely, achievement tests should not measure anything that could be labeled "ability" (Naglieri & Ford, 2005). However, the world is not structured so simply.

An extreme version of this view is the claim that "nonverbal" tests are a pure measure of abilities.

3. Things Get Complicated.

Those who understand that ability and achievement are entwined constructs confront a long list of further complications. Examples include:

The effects of culture on cognition. Beginning in the 1920s, some theorists noted that the very concept of intelligence is rooted in culture (see Anastasi, 1937; Degler, 1991), making culture-free tests of intelligence an oxymoron. Similarly, what counts as achievement varies across cultures and across time. For example, handwriting, spelling, and computation skills are valued less in the U.S. today while independent thinking and problem solving are valued more.

The effects of education, practice, and training on abilities. All abilities—from those required by the simplest reaction-time task to the most complex problem-solving task—respond to practice and training. Near the end of her career, Anastasi (1980) observed that much confusion could be avoided if "ability" were always prefaced by "developed." Similarly, R. E. Snow observed that intelligence is not only education's most important "raw material" but also its most important "product" (Snow & Yalow, 1982). This does not mean that biogical factors are unimportant. Rather, it means that all abilities are developed.

The effects of knowledge on thinking. One of the most important discoveries about human cognition is the extent to which thinking is bound to the objects of thought (Greeno, Collins, & Resnick, 1996). We now know there are no "information-free" cognitive processes. How well we reason depends on how much we know. Language has particularly powerful effects on the development of thought, from the acquisition of simple perceptual concepts to complex assemblies of knowledge and skill.

The unity of the ability/achievement space. If all abilities are achievements and all thinking is rooted in knowledge, then it makes little sense to talk about abilities and achievements as if they were qualitatively different (Snow, 1980). Rather, many who study individual differences see a single space of developed competencies or abilities (Humphreys, 1981; Cronbach, 1990; Carroll, 1993; Horn & Noll, 1997). Some develop primarily through formal schooling, others through out-of-school experiences common to most children in a given culture, and still others through experiences that are unique to the individual.

The multidimensionality of the unified ability space. For a long time we have known that ability is multidimensional. Most theorists agree that the 70+ abilities that have been identified can be organized in a hierarchy: a general "G" factor at the highest level, seven or more broad group abilities at the next level, and 50-87 primary abilities at the base (Carroll, 1993; Horn & Blankson, 2005; McGrew, 2005). *A simple division of abilities into the two camps of ability and achievement is no longer possible*. Rather, some argue that the most defensible way to view a continuum from "crystallized" achievements through "fluid" creativity is within each domain of knowledge or skill (see Snow, 1981; Sternberg, 1998).

> **If all abilities are achievements and all thinking is rooted in knowledge, it makes little sense to talk about abilities and achievements as if they were qualitatively different.**

Is Gc the real intelligence? There has long been a bias among researchers that fluid intelligence (Gf) represents the real, biologically determined intelligence, whereas crystallized intelligence (Gc) better represents the products of investing this biological intelligence in particular experiences. This view is fading among those who study abilities (Hunt, 2000). Increasingly, researchers emphasize that functional intelligence is better indexed by high levels of expertise in different domains (Ackerman, 2000; Horn & Blankson, 2005). Tests that do not measure these competencies seriously underrepresent any reasonable definition of intelligence.

The impact of affect and volition on cognition. Research on cognition shows that thinking is deeply enmeshed with affect. Interest (or disinterest), surprise (or boredom), enjoyment (or disgust) moderate what we remember about a topic, how deeply we think about it, and how long we persist in thinking about it. Further, the knowledge and skills that people assemble both reflect and feed interest. There is no way to separate the measurement of ability from motivation or feeling.

The effects of experience on brain structures. The dichotomy between biology and experience also misleads. Extensive experience in a domain effects substantial changes in the structure of the brain and the way it processes information (Nelson, 1999). This means that (1) the biological contribution to individual differences in ability is moderated by the quality of the environment and (2) experiences that enable growth determine subsequent cognitive status. Recent research on the effects of socioeconomic status on ability shows this as well (Turkheimer, Haley, Waldron, D'Onofrio, Gottesman, 2003).

The contextual specificity of thinking. Ordinarily, people who study individual differences examine only that portion of the variability in behavior that generalizes across tasks. However, what generalizes is typically only a very small portion of the variability in items. Further, factors that are specific to the test (such as the format and sample of items) are as important as the ability dimension on which the test has its highest loading. This matters for two reasons. First, test users get both parts—the portion that generalizes and the portion that is unique. This means that two different tests that load on the same factors will often give quite different scores to the same individual. Second, it means that context is more important in cognition than most ability theorists appreciate.

4. A System of Theories Is Needed.
Given the scope and complexity of research on cognition, it is not surprising that very few scholars are able to envision theories or paradigms that integrate these diverse themes into a coherent whole. Two of the most impressive efforts are those of Robert Sternberg and Richard Snow.

Sternberg has been an incredibly prolific contributor to the field. Most students are familiar with his triarchic theory of intelligence (Sternberg, 1985; 1998). The *contextual theory* addresses the abilities a particular culture values as indicants of intelligence. Sternberg argues that the major indicants of successful intelligence in contemporary U.S. culture are analytic abilities, creative abilities, and practical abilities. The *experiential theory* addresses the relative novelty of the task chosen as an indicant of intelligence. It is concerned with the effect of experience on the development of abilities. The *componential theory* considers the cognitive processes a person uses to solve a task.

Unfortunately, expelling language also expels an enormous amount of cognition, and one cannot measure the sophistication of children's reasoning or problem-solving abilities unless one presents tasks that allow children to use the knowledge they have acquired.

Snow's theory is less well known, but in many ways easier for educators to apply (for an introduction, see Corno, Cronbach, Kupermintz, Lohman, et al. 2002). The theory concerns how best to design instruction to meet the needs of different learners (Cronbach & Snow, 1977). The central construct is that of *aptitude*, by which Snow meant the degree of readiness to learn and perform well in a particular situation or domain. Aptitudes for learning tie both to what must be learned and to the learning context. Students who will have a difficult time acquiring one type of expertise (e.g., mastering algebra) may have less difficulty acquiring expertise in another domain (e.g., creative writing). Those who might have difficulty succeeding under one instructional arrangement (e.g., large lecture class) might succeed more readily under another (e.g., computer-assisted instruction).

The theory thus turns the question of "intelligence" on its head. One begins not with a catalog of the person's standing on a given set of ability dimensions but rather with a clear statement of the type of expertise one hopes to develop and of the different instructional routes available to assist learners in attaining that expertise.

Language and Thought

The use of language in testing has generated controversy largely because of (1) the linguistic diversity of children in the U.S. and (2) the erroneous belief that test scores for all children can only be interpreted by comparing them to national age or grade groups. Unfortunately, expelling language also expels an enormous amount of cognition. One cannot measure the sophistication of children's reasoning or problem-solving abilities unless one presents tasks that allow children to use the knowledge they have acquired.

Jeffrey Braden (2000) discussed this dilemma in a special issue of the *Journal of Psychoeducational Assessment* devoted to nonverbal testing. He notes that advocates of nonverbal testing rightly recognize that tests which rely on verbal knowledge may introduce construct-irrelevant variance for those students who do not speak the language on the test, who are deaf, or who have other specific linguistic disabilities. However, attempting to eliminate language and other learned symbol systems seriously restricts the construct measured by the test. *Children who speak any language and understand basic mathematical concepts cannot show how well they can reason if tests eliminate these abilities at the outset.* This is as true for minority students as it is for majority students.

Most often, for children who would do better on a nonverbal test, the problem is not the words or quantitative concepts. It is that the students' scores will be interpreted by comparing them to national norm groups. If students have lacked opportunities to develop the skills measured by an abilities test, it is important also to compare their scores to those of other students in the district who have had similar opportunities to develop these skills. This comparison group may consist of students who have had similar opportunities to learn the language of the test or may be composed of students in the same minority group (see Lohman, in press).

Measuring Reasoning Abilities

Reasoning abilities in a particular domain are a subset of a larger set of abilities that collectively index what one knows and can do in that domain. For example, verbal reasoning abilities can form part of a broad achievement construct (reading ability). A test of verbal reasoning abilities, then, would try to make salient the search, retrieval, and comparison processes that together constitute reasoning while reducing the influence of other reading processes (such as word analysis). This is why item types such as analogies are so useful. They make salient the reasoning process and, when items are well constructed, reduce the impact of other processes.

Vocabulary items on many individually administered ability tests can also measure the products of post-reasoning processes. Children infer the meanings of words by remembering and attending to similarities and differences in the way words are used in different contexts. For measuring reasoning, the critical factor is the *precision* of students' understandings of relatively common but abstract words (Snow & Lohman, 1989).

The Sentence Completion test on *CogAT* also emphasizes reasoning. In a recent article, Naglieri and Ford (2005) focused on this subtest as an example of a task that "blurs the line between ability and achievement." They looked at the distribution of "readability" for the twenty items at Level D, using the Flesch-Kincaid method which combines the number of words and syllables in a passage to predict how difficult it might be to read. The values for different sentences varied widely; the average grade level score was 6.1, a value they called alarming. However, their data have no merit.

Those with expertise in the construction of such tests have long known (Oakland & Lane, 2004) not to use readability formulas unless passages have at least one hundred words. Using readability formulas on individual sentences essentially produces a random variable. If reading items were a problem, then the difficulties of items should be predicted by their readability. However, item difficulty on the Sentence Completion test and estimates of readability are not significantly correlated (at some test levels the correlation is negative). In other words, there is *no* evidence that students miss Sentence Completion items because they cannot read them.

Finally, from an aptitude perspective, one wants to know how well students can perform cognitive operations that (collectively) demonstrate verbal reasoning. Students who have difficulty doing this on tests that use fairly common words and short sentences will also have difficulty understanding long prose passages or less-structured teacher explanations. Not measuring these abilities does not make them any less relevant for learning or performance in the classroom.

Implications for Identification and Instruction

CogAT is frequently used to help in the process of identifying (1) students who currently display extraordinary levels of accomplishment in domains for which accelerated instruction is offered and (2) students who show promise for developing such academic excellence. Educators must match the students' aptitudes to the instruction offered. Children who excel in writing narratives may not excel in mathematical reasoning. Children who can succeed when working with others may not succeed when working alone (Snow & Lohman, 1984).

Understanding readiness (or aptitude) is generally much more helpful than trying to measure intelligence. What must students know and be able to do in order to attain expertise? Research on academic learning shows that the best predictors of subsequent learning in a domain are (1) current achievement in that subject area, (2) the ability to reason in the symbol systems of that domain, (3) interest in that subject area, and (4) the willingness to persist in order to attain excellence (Corno et al. 2002). [*Note:* The winter 2005 issue of *Cognitively Speaking* details how to identify academically gifted minority students using this aptitude approach. The winter issue is available online at www.cogat.com.]

This aptitude approach is especially helpful in identifying academically talented minority students. For example, one can usefully ask whether an ELL student displays sufficient readiness to learn in a classroom in which English is the language of instruction. If English proficiency is low, then one can intervene to improve those skills in reading and reasoning with English that are critical aptitudes for classroom learning.

Understanding that all abilities are grounded in biological processes and yet are developed has broad implications for programs for the academically gifted. The best accelerated programs see their mission as developing talent—not merely discovering it. Educators can emphasize their role in developing academic excellence. All students can strive to achieve excellence.

Conclusions

Ability tests are perhaps best understood as achievement tests of a special sort. Conversely, achievement tests may be seen as ability tests of a special sort. The path from a belief that ability and achievement are separate to the inclusive systems theories of Sternberg (1998) or Snow (1994) has many stopping-off points. A willingness to move beyond one's current beliefs has always characterized more thoughtful professionals. As in other areas of science, one of the first steps in this process is to articulate one's beliefs. The second is to look closely at the disconcerting evidence that challenges rather than reinforces intuitions. The first goal of this paper is to provide a few such challenges. The second goal is to describe something of the rewards of the journey ahead for those willing to embark on it.

References

Ackerman, P. L. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence: Gf/Gc, personality and interest correlates. *The Journals of Gerontology: Psychological Sciences, 55B (2)*, 69-84.

Anastasi, A. (1937). *Differential psychology*. New York: Macmillan.

Anastasi, A. (1980). Abilities and the measurement of achievement (pp. 1-10). In W. B. Schrader (Ed.), *Measuring achievement: Progress over a decade: Proceedings of the 1979 ETS Invitational Conference.* San Francisco: Jossey Bass. *(New directions for testing and measurement series).*

Braden, J. P. (2000). Perspectives on the nonverbal assessment of intelligence. *Journal of Psychoeducational Assessment, 18*, 204-210.

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

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153-193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1-22.

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

Corno, L., Cronbach, L. J., Kupermintz, H., Lohman, D. F., et al. (2002). *Remaking the concept of aptitude: Extending the legacy of Richard E. Snow.* Hillsdale, NJ: Erlbaum.

Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper & Row.

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

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

Degler, C. (1991). *In search of human nature: The decline and revival of Darwinism in American social thought.* New York: Oxford University Press.

Ferguson, G. A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology, 10,* 121-131.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.

Horn, J. L. & Blankson, N. (2005). Foundations for better understanding cognitive abilities. In Flanagan, D. P. & Harrison, P. L. (Eds.), *Contemporary intellectual assessment: Theories, tests, & issues (2nd ed.)* (pp. 41-68). NY: The Guilford Press.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-91). New York: Guilford.

Humphreys, L. G. (1981). The primary mental ability. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 87-102). New York: Plenum.

Hunt, E. B. (2000). Let's hear it for crystallized intelligence. *Learning and Individual Differences, 12,* 123-130.

Jensen, A. R. (1998) *The g factor: The science of mental ability.* Westport, CT: Praeger.

Lohman, D. F. (1993). Teaching and testing to develop fluid abilities. *Educational Researcher, 22*, 12-23.

Lohman, D. F. (2005). The role of nonverbal ability tests in the identification of academically gifted students: An aptitude perspective. *Gifted Child Quarterly, 49,* 111-138.

Lohman, D. F. (in press). An aptitude perspective on talent identification: Implications for the identification of academically gifted minority students. *Journal for the Education of the Gifted.*

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues (2nd ed.)* (pp. 136-182). New York: The Guilford Press.

Naglieri, J. A., & Ford, D. (2005). Increasing minority children's participation in gifted classes using the NNAT: A response to Lohman. *Gifted Child Quarterly, 49*, 29-36.

Nelson, C. A. (1999). Neural plasticity and human development. *Current Directions in Psychological Science, 8,* 42-45.

Oakland, T., & Lane, H. B. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing, 4,* 239-252.

Snow, R. E. (1980). Aptitude and achievement (pp. 39-60). In W. B. Schrader (Ed.) *New directions of testing and measurement; Measuring achievement: Progress over a decade.* San Francisco: Jossey Bass.

Snow, R. E. (1981). Toward a theory of aptitude for learning: Fluid and crystallized abilities and their correlates. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 345-362), New York: Plenum.

Snow, R. E. (1994). Abilities in academic tasks. In R. J. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 3-37). Cambridge, UK: Cambridge University Press.

Snow, R. E., & Lohman, D. F. (1984). Toward a theory of aptitude for learning from instruction. *Journal of Educational Psychology, 76*, 347-376.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263-331). New York: Macmillan.

Snow, R. E., & Yalow, E. (1982). Education and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 493-585). Cambridge, UK: Cambridge University Press.

Spearman, C. E. (1923). *The nature of intelligence and the principles of cognition.* London: Macmillan.

Stern, W. (1914). The psychological methods of testing intelligence *(Educational Psychology Monographs No. 13).* Baltimore: Warwick & York. (G. M. Whipple, trans.)

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

Sternberg, R. J. (1998). Applying the triarchic theory of human intelligence in the classroom. In R. J. Sternberg & W. M. Williams (Eds.), *Intelligence, instruction, and assessment: Theory into practice.* Mahwah, NJ: Erlbaum.

Thorndike, E. L., Bregman, E. O., Cobb, M. V., & Woodyard, E. (1926). *The measurement of intelligence.* New York: Columbia University, Teachers College, Bureau of Publications.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14,* 623-628.

Riverside Publishing
A HOUGHTON MIFFLIN COMPANY

Cognitively Speaking **is a publication of Riverside Publishing.**

Riverside Publishing
425 Spring Lake Drive
Itasca, IL 60143-2079

For more information, call Customer Service at 800.323.9540 or visit www.riversidepublishing.com and www.cogat.com.