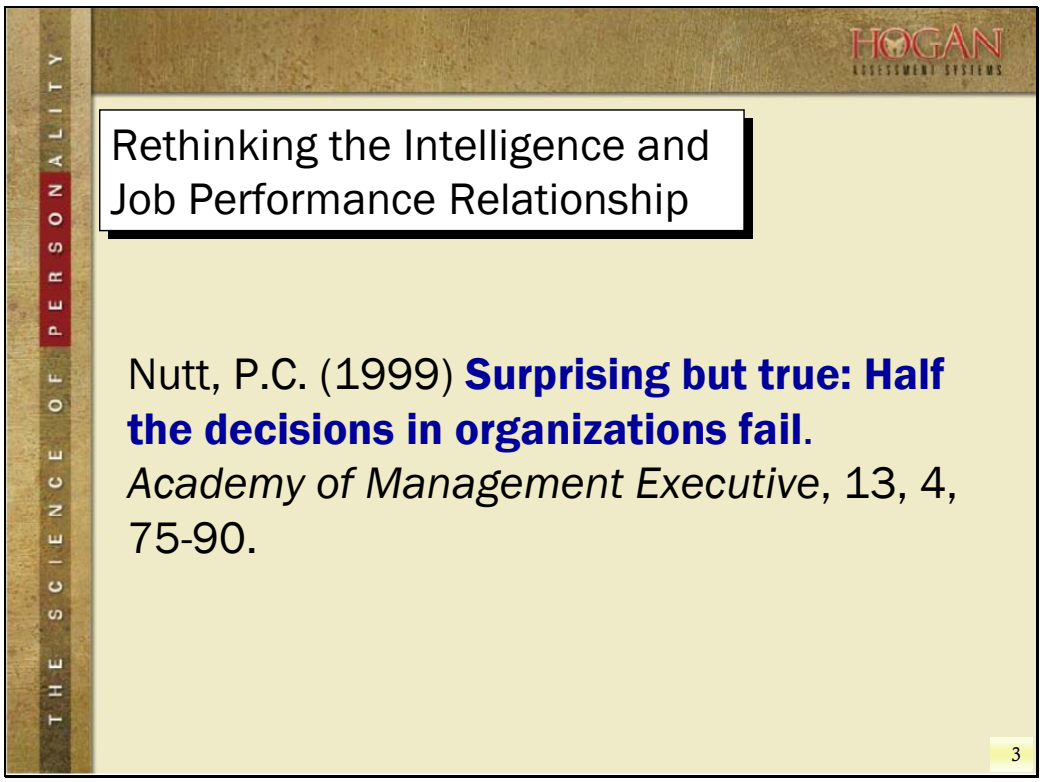
THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Good Judgment

The Intersection of Intelligence and Personality

Robert Hogan & **Paul Barrett**
President Chief Research Scientist



The Abstract to the paper

Half the decisions in organizations fail. Studies of 356 decisions in medium to large organizations in the U.S. and Canada reveal that these failures can be traced to managers who impose solutions, limit the search for alternatives, and use power to implement their plans. Managers who make the need for action clear at the outset, set objectives, carry out an unrestricted search for solutions, and get key people to participate are more apt to be successful. Tactics prone to fail were used in two of every three decisions that were studied.

From the first few paragraphs of the paper

Why Decisions Fail

Half the decisions made in organizations fail. This is a major finding from studies of decision making that I have conducted over the past two decades. This dramatic failure rate cannot be explained by conventional wisdom. For instance, failure does not generally stem from things beyond a manager's control: draconian regulations imposed by government, unexpected budget cuts by higher-ups, or loss of market share because of fickle customers.

Although failures can occur when regulations run up costs, when budget flexibility is lost, and when customer preferences shift and wreck a marketing plan, the tactics that managers use are more important. Studies of decision making that I have conducted over a 20-year period trace failure to managers who employ poor tactics.

Why are tactics prone to fail used so often? My work suggests three reasons: Some tactics with a good track record are commonly known, but uncommonly practiced. Nearly everyone surveyed is aware of participation and its ability to coax acceptance, but participation is used in just one of five decisions.

Another reason is that decision makers take short cuts when they feel time pressure. As pressure appears to mount, reason gives way to such homilies as, "Why invent the wheel?" The practices of a respected organization are then copied, which is rationalized as timely and pragmatic. But such short

cuts often lead to unanticipated difficulties and delays as steps are taken to convince people that the company's interests, not yours, are being served.

A third reason for failure is subtleties. Telling people you want to lower costs is much more powerful than finding the root cause of the cost problem. Managers who are drawn to problem solving fail to see that problems prompt blame. Telling subordinates what's wanted liberates them to look for answers. Finding problems alerts subordinates to take defensive action. Energy is directed away from finding answers to protecting their backs and their interests."

The results of this paper indicates that higher ability (as is the case for the majority of managers who are required to have undertaken tertiary education and acquired at least an undergraduate degree) does not automatically lead to superior decision-making performance. Yet, surely higher intelligence should mean better decision-making and better judgments?

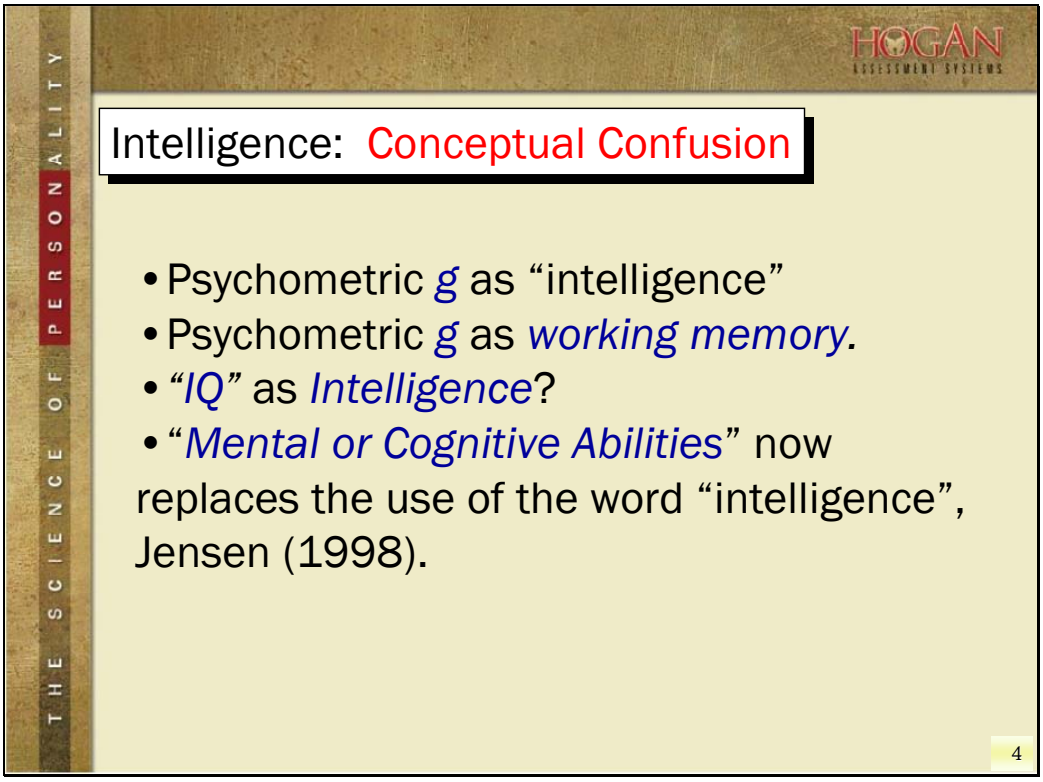
Part of the reason why businesses and organizations are successful is because their leaders and managers make good judgments when called upon to make decisions.

Can this "Good Judgment" be measured? Is it synonymous with "intelligence"?

To us, there seems to be two issues here:

1. What do we actually mean by intelligence.
2. Is "good judgment" a combination of both intelligence and personality?

When we look at the "psychometrics" of intelligence – we see conceptual confusion…



Intelligence: a meaningful construct or mere distraction?

Jensen, A.R. (1998) *The g Factor: The Science of Mental Ability*. Praeger, p.48, 3rd – and onwards, provides probably the best "lead in" to recommending the discarding of the use of the word "intelligence" altogether …

> *"My study of these two symposia and of many other equally serious attempts to define "intelligence" in purely verbal terms has convinced me that psychologists are incapable of reaching a consensus on its definition. It has proved to be a hopeless quest.* ***Therefore, the term "intelligence" should be discarded altogether in scientific psychology, just as it discarded "animal magnetism" as the science of chemistry discarded "phlogiston." "Intelligence" will continue, of course, in popular parlance and in literary usage, where it may serve a purpose only because it can mean anything the user intends, and where a precise and operational definition is not important.***
>
> *Largely because of its popular and literary usage, the word "intelligence" has come to mean too many different things to many people (including psychologists). It has also become so fraught with value judgments, emotions, and prejudices as to render it useless in scientific discussion.[" I have no quarrel with the typical dictionary definition of "intelligence," except that it does not adequately describe what I am actually writing about in this book. Indeed the attempt to provide a purely lexical definition is a hindrance to understanding the phenomena examined here.*
>
> *I am certainly not proposing to offer still another definition of "intelligence" or another term to take its place, or suggesting that any existing definition can be made more acceptable if we modify it with adjectives, such as "academic intelligence" and "practical intelligence," or by making it plural, such as the currently popular "multiple intelligences." These "solutions" can only worsen the problem. As an intraspecies concept it is best simply to get rid of the term "intelligence" altogether in psychology. In this we should follow Spearman and henceforth drop the ill-fated word from our scientific vocabulary, or use it only in quotes, to remind ourselves that it is not only scientifically unsatisfactory but wholly unnecessary.*
>
> *Formal definitions, however, are essential in science. But they themselves do not address the nature of a phenomenon or claim the status of empirical fact. Nor do they imply any particular theory or require logical or experimental proof. Formal definitions are theoretically neutral conventions that scientists agree upon in order to get on with their job. It makes no sense to disagree over such conventional definitions. It is important only that a definition be sufficiently clear and explicit to serve its pragmatic purpose.*
>
> *To put the study of mental ability on a firm scientific footing, we must begin by using theoretically neutral, objective, operational definitions. From this position "intelligence" (or any synonym or conceptual substitute for it) never needs to enter the discussion. Just blot out whatever this word with all its ambiguities and emotional baggage may mean to you (or your dictionary). (Nor shall I try to provide another word or concept as a verbal substitute.)"*

From then on, Jensen uses the word "ability" or phrase "mental ability". The critical distinction here is between the technical/specific concept of a mental ability as espoused by Jensen, such as 3-D mental rotation, verbal comprehension, digit-span forward, picture completion, ideational fluency etc., and the two meta-concepts of "think clearly" and "good judgment".

Only recently, Hunt, E., & Carlson, J. (2007a, on page 199) noted that :

Just as there have been disputes about the reality and nature of race, there have been disputes about the reality and nature of intelligence. After reviewing several discussions of this topic, Browne-Miller (1995, p. 37) noted that the following opinions have been expressed (which we summarize slightly to save space):

(1) Intelligence does not exist.
(2) Intelligence does exist and is measured by intelligence tests.
(3) Intelligence does exist and is not measured by intelligence tests.
(4) Intelligence is unchangeable, fixed at birth.
(5) Intelligence is either partially or entirely environmentally determined; as a corollary, intelligence is learnable, especially through education.
(6) Intelligence is something that is the same as, or overlaps with, learning ability.
(7) Intelligence is purely cognitive.
(8) Intelligence can take many forms, in domains as diverse asmusic, mathematics, athletics, and leadership.

Evidently intelligence is at least as fuzzy a concept as race is!

Our own view is a mixture of points (2), (5), (6), and (7). We believe that intelligence exists; that it is partially measured by intelligence tests; that it is both environmentally and genetically determined; that it is closely related to, but not quite the same as, learning ability;4 and that extending the term intelligence beyond the cognitive area so expands the domain that it becomes an unmanageable concept.

But, note Sternberg and Grigorenko's response in the same issue … (p. 222)

The second unsupported argument in the Hunt and Carlson article is that ''extending the term intelligence beyond the cognitive area so expands the domain that it becomes an unmanageable concept'' (p. 199). This statement, like the previous statement about race, reflects the authors' socialized level of comfort rather than any scientific fact or principle

In their response to their commentators, Hunt and Carlson (2007b, p. 224) state …

We said that so many definitions have been given that intelligence itself is a "fuzzy" (i.e., ill-defined) concept. Gottfredson claims this is not so, and offers what she presents as an agreed-upon definition, based on statistical analyses of intelligence test scores. Sternberg and Grigorenko take us to task for not including emotionality and personal goals, which are part of Sternberg's notion of successful intelligence. Considering that these opposing criticisms come from well-established, knowledgeable investigators, we maintain our point. The definition of intelligence is fuzzy.

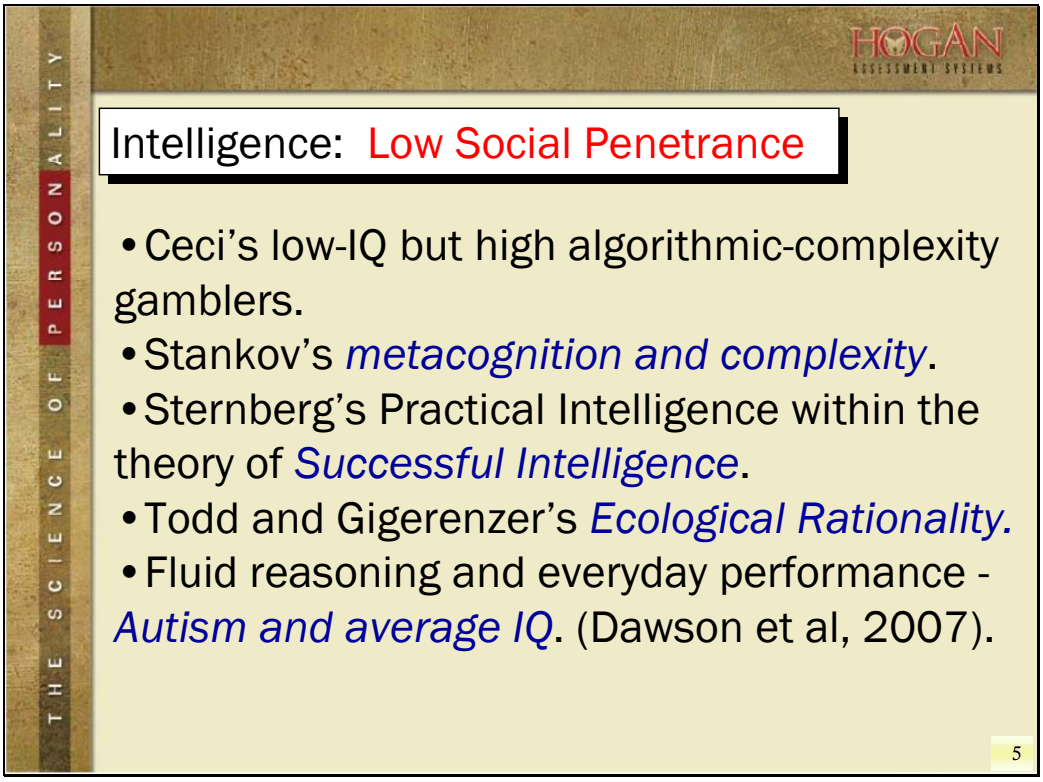
One of the important steps in science is the creation of definitions that, in Plato's terms, "carve nature at the joints." The distinction between acceleration and velocity was crucial to the advancement of physics. The study of intelligence would benefit by an analysis of clear definitions rather than claims that one or another group of investigators has discovered the right one. We believe that Gottfredson's emphasis on extant tests is too narrow, but that Sternberg and Grigorenko would cast such a broad net that it would make the study of intelligence unmanageable. We offered an intermediate position. We hope this matter will receive discussion in future papers.

The problem for Jensen, Hunt, Carlson, and others is that an entire field of Artificial Intelligence (or Computational Intelligence) exists around a clear definition of that word "Intelligence". Indeed, I think it makes no sense at all to treat the word "Intelligence" as some "non-word", on the basis of an argument about a plethora of definitions and the lack of specificity of what the construct might mean. It is one thing

to ask for technical definitions of specific attributes like "mental abilities" – quite another to deny the value of a concept which, although very broad in meaning, does seem to be used quite readily by the rest of humanity (*although not by psychometricians*) and scientists who actually attempt to create and build intelligent machines.

However, another problem with this psychometric view of "intelligence", IQ, 'g', or specific mental abilities is that it seems to have little relation to everyday performance of human beings. This does not deny its value in predicting educational outcomes, occupational levels of jobs, or indeed annual income, but it seems strangely deficient when called upon to predict what might be called everyday performance – where groups of people with low IQs or low psychometric 'g' seem nevertheless to be using sophisticated decision algorithms or learned behaviors which should have been impossible to acquire given the assumed cognitive inefficiencies of a low IQ.

Usually, these phenomena are written off by differential psychologists as "domain-specific" skills or highly specialized skills which only work in specific contexts. The problem though is how a supposedly cognitively deficient brain (poor working memory, poor information processing, and poor executive functions) can acquire what seem to be fairly sophisticated decision-making skills).



Likewise, with all the correlational evidence associated with brain size, grey and white matter densities, and IQ", how is that an individual with nearly 75% of his brain "missing" can still have an IQ of 84 on verbal ability, 75 full-scale, and apparently function quite normally in life?

Fuillet, Dufour, & Pelletier (The Lancet, July 2007) report the case of a 44 year old man with massive ventricular enlargement (hydrocephalus) …

> *A 44-year-old man presented with a 2-week history of mild left leg weakness. At the age of 6 months, he had undergone a ventriculoatrial shunt, because of postnatal hydrocephalus of unknown cause. When he was 14 years old, he developed ataxia and paresis of the left leg,*

which resolved entirely after shunt revision. His neurological development and medical history were otherwise normal. He was a married father of two children, and worked as a civil servant. On neuropsychological testing, he proved to have an intelligence quotient (IQ) of 75: his verbal IQ was 84, and his performance IQ 70. CT showed severe dilatation of the lateral ventricles (see the figure below); MRI revealed massive enlargement of the lateral, third, and fourth ventricles, a very thin cortical mantle and a posterior fossa cyst.

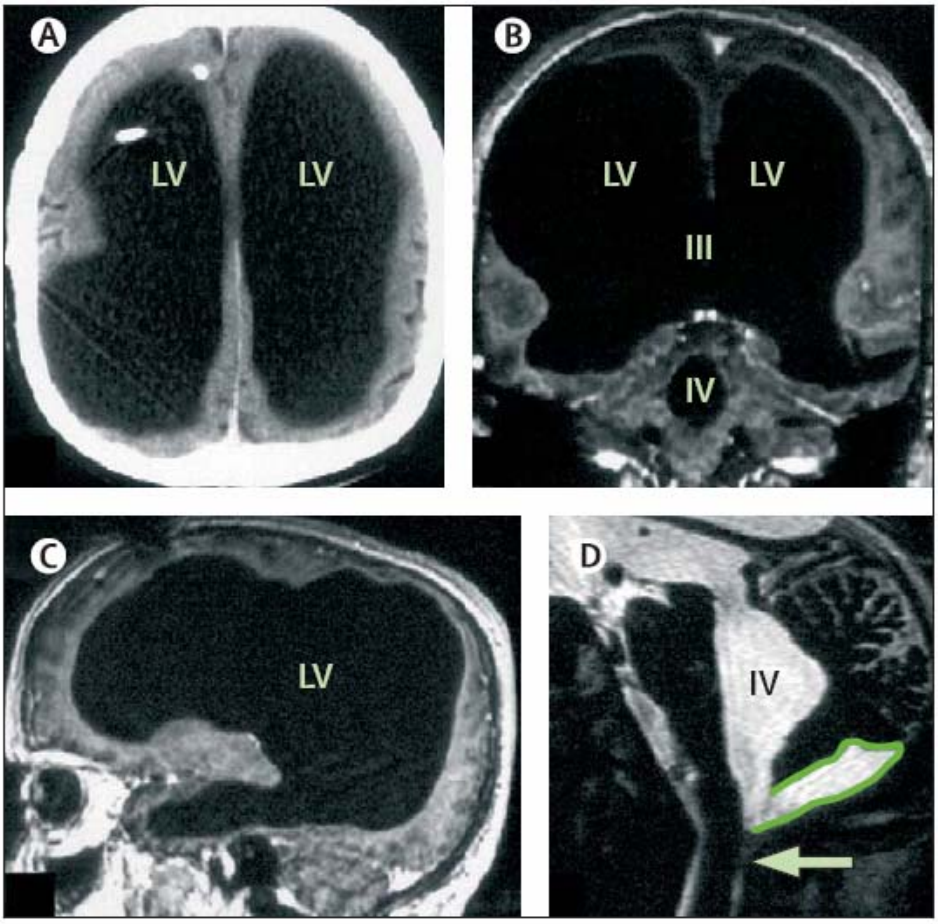


Figure: **Massive ventricular enlargement, in a patient with normal social functioning**
(A) CT; (B, C) T1- weighted MRI, with gadolinium contrast; (D) T2-weighted MRI. LV=lateral ventricle. III=third ventricle. IV=fourth ventricle. Arrow=Magendie's foramen. The posterior fossa cyst is outlined in (D).

From a report in the New Scientist magazine, 20th July, 2007 …

A man with an unusually tiny brain manages to live an entirely normal life despite his condition, which was caused by a fluid build-up in his skull. Scans of the 44-year-old man's brain showed that a huge fluid-filled chamber called a ventricle took up most of the room in his skull, leaving little more than a thin sheet of actual brain tissue (see image below).

"It is hard for me [to say] exactly the percentage of reduction of the brain, since we did not use software to measure its volume. But visually, it is more than a 50% to 75% reduction," says Lionel Feuillet, a neurologist at the Mediterranean University in Marseille, France.

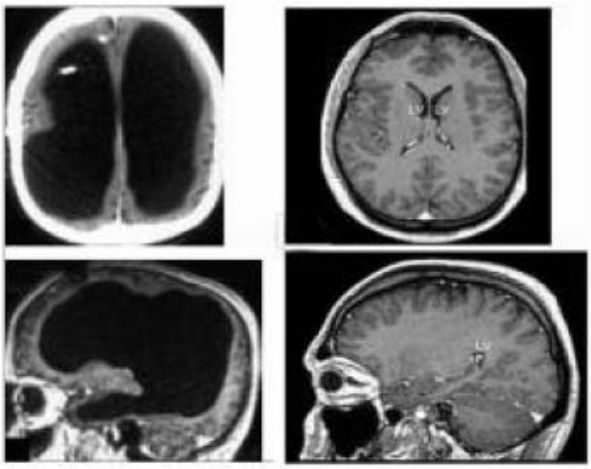
Feuillet and his colleagues describe the case of this patient in The Lancet. He is a married father of two children, and works as a civil servant.

Not retarded

The man went to a hospital after he had mild weakness in his left leg. When Feuillet's staff took his medical history, they learned that, as an infant, he had had a shunt inserted into his head to drain away hydrocephalus – water on the brain. The shunt was removed when he was 14. But the researchers decided to check the condition of his brain using computed tomography (CT) scanning technology and another type of scan called magnetic resonance imaging (MRI). They were astonished to see "massive enlargement" of the lateral ventricles – usually tiny chambers that hold the cerebrospinal fluid that cushions the brain. Intelligence tests showed the man had an IQ of 75, below the average score of 100 but not considered mentally retarded or disabled. "The whole brain was reduced – frontal, parietal, temporal and occipital lobes – on both left and right sides. These regions control motion, sensibility, language, vision, audition, and emotional and cognitive functions," Feuillet told New Scientist.

Brain adaptation

The findings reveal "the brain is very plastic and can adapt to some brain damage occurring in the preand postnatal period when treated appropriately," he says.



Enlarge image

The large black space shows the fluid that replaced much of the patient's brain (left). For comparison, the images (right) show a typical brain without any abnormalities (Images: Feuillet et al./*The Lancet*)

Clearly, while such cases are truly exceptional, it only takes one case such as this to show that neural tissue is considerably more adaptive than most differential psychologists appear to argue. As to the behavior genetics arguments which propose a considerable biologically-fixed processing capacity of individuals – well, given the latest news on microRNA, even this behavior-genetic work from psychologists is now beginning to look problematic.

A briefing paper in the Economist magazine (June 16th, 2007) entitled "Really New Advances" begins:

IT IS beginning to dawn on biologists that they may have got it wrong. Not completely wrong, but wrong enough to be embarrassing. For half a century their subject had been built around the

relation between two sorts of chemical. Proteins, in the form of enzymes, hormones and so on, made things happen. DNA, in the form of genes, contained the instructions for making proteins. Other molecules were involved, of course. Sugars and fats were abundant (too abundant, in some people). And various vitamins and minerals made an appearance, as well. Oh, and there was also a curious chemical called RNA, which looked a bit like DNA but wasn't. It obediently carried genetic information from DNA in the nucleus to the places in the cell where proteins are made, rounded up the amino-acid units out of which those proteins are constructed, and was found in the protein factories themselves.

All that was worked out decades ago. Since then, RNA has been more or less neglected as a humble carrier of messages and fetcher of building materials. This account of the cell was so satisfying to biologists that few bothered to look beyond it. But they are looking now. For, suddenly, cells seem to be full of RNA doing who-knows-what.

And the diversity is staggering. There are scnRNAs, snRNAs and snoRNAs. There are rasiRNAs, tasiRNAs and natsiRNAs. The piRNAs, which were discovered last summer, are abundant in developing sex cells. No male mammal, nor male fish, nor fly of either sex, would be fertile without them. Another RNA, called XIST, has the power to turn off an entire chromosome. It does so in females because they, unlike males, have two X chromosomes and would otherwise get an unhealthy double dose of many proteins. There is even a "pregnancy-induced non-coding RNA", cutely termed PINC. New RNAs are rushing forth from laboratories so rapidly that a group called the RNA Ontology Consortium has been promised half a million dollars to prune and tend the growing thicket of RNA-tailed acronyms.

In the light of this abundance, perceptions about what a gene is need to change. Genes were once thought of almost exclusively as repositories of information about how to build proteins. Now, they need to be seen for what they really are: RNA factories. Genes for proteins may even be in the minority. In a human, the number of different microRNAs, one of the commonest of the newly discovered sorts of RNA, may be as high as 37,000 according to Isidore Rigoutsos, IBM's genome-miner in chief. That compares with the 21,000 or so protein-encoding genes that people have.

Philosophers of science love this sort of thing. They refer to it as a paradigm shift. Living through such a shift is confusing for the scientists involved, and this one is no exception. But when it is over, it is likely to have changed people's views about how cells regulate themselves, how life becomes more complex, how certain mysterious diseases develop and even how the process of evolution operates. As a bonus, it also opens up avenues to develop new drugs.

And further down ...

Ronald Plasterk, of the University of Utrecht, in the Netherlands, suggests that microRNAs are important in the evolution of the human brain. In December's Nature Genetics, he compared the microRNAs encoded by chimpanzee and human genomes. About 8% of the microRNAs that are expressed in the human brain were unique to it, much more than chance and the evolutionary distance between chimps and people would predict.

Such observations suggest evolution is as much about changes in the genes for small RNAs as in the genes for proteins—and in complex creatures possibly more so. Indeed, some researchers go further. They suggest that RNA could itself provide an alternative evolutionary substrate. That is because RNA sometimes carries genetic information down the generations independently of DNA, by hitching a lift in the sex cells. Link this with the fact that the expression of RNA is, in certain circumstances, governed by environmental factors, and some very murky waters are stirred up.

What is being proposed is the inheritance of characteristics acquired during an individual's lifetime, rather than as the result of chance mutations. *This was first suggested by Jean Baptiste Lamarck, before Charles Darwin's idea of natural selection swept the board. However, even Darwin did not reject the idea that Lamarckian inheritance had some part to play, and it did not disappear as a serious idea until 20th-century genetic experiments failed to find evidence for it.*

The wiggle room for the re-admission of Lamarck's ideas comes from the discovery that small RNAs are active in cells' nuclei as well as in their outer reaches. Greg Hannon, of the Cold Spring Harbor Laboratory in New York State, thinks that some of these RNA molecules are helping to direct subtle chemical modifications to DNA. Such modifications make it harder for a cell's code-reading machinery to get at the affected region of the genome. They thus change the effective composition of the genome in a way similar
microRNA – to mutation of the DNA itself (it is such mutations that are the raw material of natural selection). Indeed, they sometimes stimulate actual chemical changes in the DNA—in other words, real mutations.

Even this observation, interesting though it is, does not restore Lamarckism because such changes are not necessarily advantageous. But what Dr Hannon believes is that the changes in question sometimes happen in response to stimuli in the environment. The chances are that even this is still a random process, and that offspring born with such environmentally induced changes are no more likely to benefit than if those changes had been induced by a chemical or a dose of radiation. And yet, it is just possible Dr Hannon is on to something. The idea that the RNA operating system which is emerging into view can, as it were, re-write the DNA hard-drive in a predesigned way, is not completely ridiculous.

This could not result in genuine novelty. That must still come from natural selection. But it might optimise the next generation using the experience of the present one, even though the optimising software is the result of Darwinism. And if that turned out to be commonplace, it would be the paradigm shift to end them all.

Now, here we are way beyond the simple-minded rhetoric of the “race realists” and gene-centric behavior-genetic psychologists. It also explains why years of searching for “genes for intelligence” have been in vain. The door to the new thinking and theorizing about human intelligence is now open.

The recent paper by Todd and Gigerenzer (2007) on ecological rationality also shows how some psychological scientists are now beginning to finally move beyond the old “psychometric-statistical” trait models to address more closely the factual observations from the real world.

The importance of looking at the world to understand the mind has long been appreciated by a few prominent thinkers. Charles Darwin held that environmental forces had shaped human behavior through natural selection, leading to the modern call by evolutionary psychologists to look to our ancestral world for the problems ourminds are designed to solve.More than 50 years ago, Egon Brunswik urged psychologists to study the texture of natural environments and the corresponding structure of cues the mind relies on to infer the state of its surroundings. Roger Shepard spoke of the mind as a mirror, reflecting long-standing physical aspects of the world such as the 24-hour light–dark cycle. Herbert Simon proposed the metaphor of the mind and world fitting together like the blades of a pair of scissors—the two must be well matched for effective behavior to be produced, and just looking at the cognitive blade will not explain how the scissors cut. In each case, the world is a key for understanding the workings of the mind (Todd & Gigerenzer, 2001). However, prevailing explanations of behavior are still expressed most often in terms of personality traits, cognitive styles, brain-region activation patterns, preferences and utilities, and other assumed entities ‘‘inside’’ the mind. (p. 167)

Indeed, Lazar Stankov's (2003) thoughtful chapter entitled "Complexity in Human Intelligence" has a nice statement on page 40 which seems to summarize why we have chosen to attempt to measure "Good Judgment" rather than just create yet another "Critical Reasoning" or "mental ability" test…

> *Furthermore, aspects of metacognition extend into the areas bordering on personality, including cognitive planning, monitoring of performance, evaluation, and so on. Although these processes are often considered as being distinct from intelligence, they cannot be ignored in truly comprehensive accounts of human cognition.*

Before we depart this section, it is also worth noting that other domains of psychology are also rebelling against the simplistic "quantitative imperative" or more accurately "statistical imperative" imposed by psychometricians and psychologists who seem unable to comprehend that human beings are not physical systems whose cause-effect relations can be modeled as though they were examining physical objects as in the physical sciences …

Take, for example, the careful and lengthy examination of the field of Developmental Psychopathology by John Richters (1997), the abstract is given below …

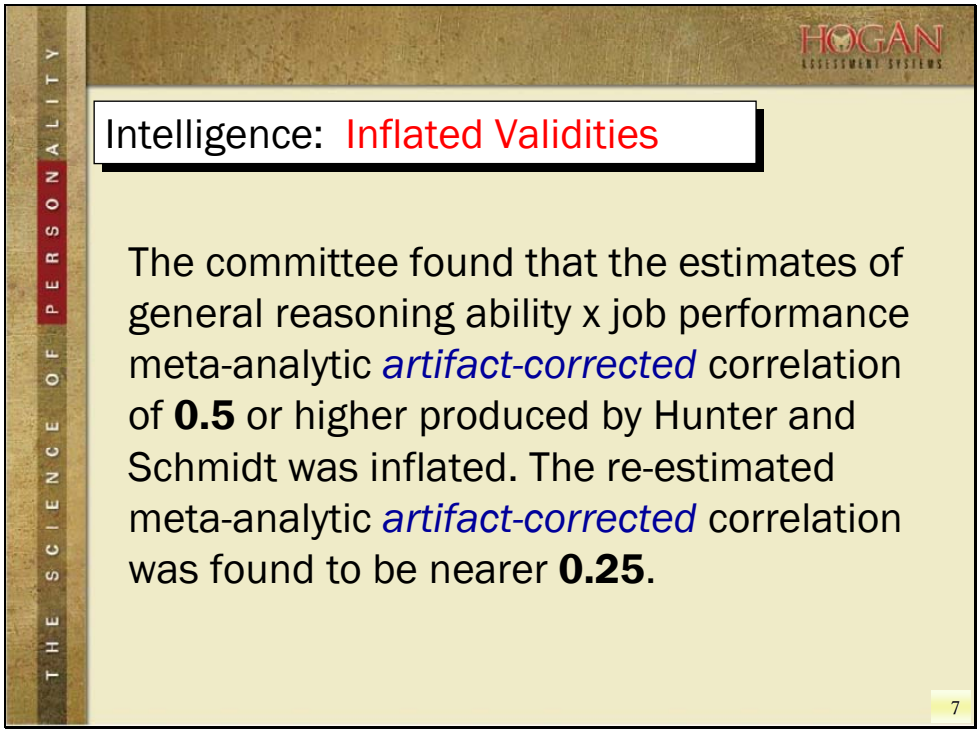
> *Developmental psychopathology stands poised at the close of the 20th century on the horns of a major scientific dilemma. The essence of this dilemma lies in the contrast between its heuristically rich open system concepts on the one hand, and the closed system paradigm it adopted from mainstream psychology for investigating those models on the other. Many of the research methods, assessment strategies, and data analytic models of psychology's paradigm are predicated on closed system assumptions and explanatory models. Thus, they are fundamentally inadequate for studying humans, who are unparalleled among open systems in their wide ranging capacities for equifinal and multifinal functioning. Developmental psychopathology faces two challenges in successfully negotiating the developmentalist's dilemma. The first lies in recognizing how the current paradigm encourages research practices that are antithetical to developmental principles, yet continue to flourish. I argue that the developmentalist's dilemma is sustained by long standing, mutually enabling weaknesses in the paradigm's discovery methods and scientific standards. These interdependent weaknesses function like a distorted lens on the research process by variously sustaining the illusion of theoretical progress, obscuring the need for fundamental reforms, and both constraining and misguiding reform efforts. An understanding of how these influences arise and take their toll provides a foundation and rationale for engaging the second challenge. The essence of this challenge will be finding ways to resolve the developmentalist's dilemma outside the constraints of the existing paradigm by developing indigenous research strategies, methods, and standards with fidelity to the complexity of developmental phenomena.*

We hope the reader can now begin to appreciate better why are striving to generate a new kind of psychological assessment which goes beyond the taxonomies of "mental ability" psychology and begins to make assessments of clearly important and useful constructs like "Good Judgment". This requires innovation beyond simply tweaking test technology or a few items.

The concept itself is fuzzy (in a technical sense) – but, it is not immune to empirical examination and empirical evidence accrual. But, the methods for acquiring such evidence will not necessarily rely upon the routine application of multivariate statistics and "rubber-stamp" significance tests.

Instead, as scientists, we will fit the analysis methods to the theory and problem at hand. We are aiming to design and measure a construct which is recognized as important by business and organizations, and not to impose upon them assessments which measure constructs which happen to be "tractable" to a certain restricted class of linear statistical methods and assumptions.

Slides 6-11 go on to note that the body of evidence upon which many I/O psychologists rely to give validity to the statement "IQ tests/GMA substantially predict Job Performance" may itself be problematic. The evidence for this is drawn from two sources – a published investigative research report from the US National research Council, and some simple data simulations showing the "spin" vs the "reality" of working with "corrected" validity coefficients.



Let's look at what Henry Levin (1991) p. 358-359, had to say of the investigative report …

Hartigan, J.A., and Wigdor, A.K. (1989) Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery. National Academy Press.

> *It is well known that racial minorities tend to score considerably lower on standardized test scores than Whites. If tests were used to select workers from an applicant pool, straightforward ranking of applicants by their test scores would have a large adverse impact on minorities. For example, Blacks typically score about a standard deviation below Whites on tests of verbal and mathematical proficiency. But, the fact that such tests predict measures of worker performance so imperfectly means that they cannot be the sole criterion of employee selection if fairness and efficiency are the goals.*
>
> *This is the setting for this important volume. The U.S. Employment Service screens some 19 million job applicants a year through its related state job service agencies. About 400 of the 1,800 local affiliates use the General Aptitude Test Battery (GATB) for employee referrals. However, all scores are converted to within-group percentiles for Blacks, Hispanics, and others (non- Black and non-Hispanic candidates). Job applicants are recommended to employers on the basis of a top-down ranking system where a candidate with a given percentile ranking within a racial group is placed in the same bracket of preferment as a candidate with the identical percentile rank from the other racial groups. Thus, a Black or Hispanic in the 78th percentile would be referred over a White in the 77th percentile, even though the raw score on GATB for the White would be considerably higher.*
>
> *This practice has been viewed by the U.S. Department of Labor as supporting both the enhancement of economic productivity through top-down referral and the equalization of employment opportunity for minorities through within-group scoring. In the middle 1980s, the Civil Rights branch of the U.S. Department of Justice argued that the within-group scoring procedure must be abandoned because it discriminates against White job candidates. The Department of Labor argued that it had adopted such a policy to prevent discrimination against racial minorities. This skirmish between the two agencies led to a stalemate in which it was agreed that the present system of GATB referral would neither expand nor be legally attacked until a study of the referral system could be carried out by a panel of experts. This volume is a response to that request.*
>
> *It reports the results of a study of a panel of experts of the National Research Council, chaired by John A. Hartigan of the Department of Statistics at Yale. The panel comprised specialists in employment testing, industrial psychologists, economists, and a number of noted specialists in educational measurement including Richard Jaeger, Robert Linn, Lorrie Shepard, and Larry Hedges.*

The panel concluded (p. 360):

> *Perhaps the most interesting part of the report is found in the evaluations of corrected validities. The Employment Service has estimated that the true validities are .5 or higher, whereas the Committee found that they were about half of this value. The differences stem from two major divergences. First, the Employment Service relied on the results from some 515 validity studies performed prior to 1972 that showed observed correlations of about .25 between cognitive, perceptual, or psychomotor aptitude scores and supervisor ratings on the job. However, this observed correlation falls to about .19 in some 264 studies done since 1972. The Committee placed more credence in the later studies than the earlier ones.*
>
> *Second, the Committee concluded that the corrections used by Hunter and Schmidt (1982) and the Employment Service vastly overstated the true validities. The Committee started with the*

average observed validity of GATB aptitude composites on supervisor ratings of about .22 for the five job families, based on the post-1972 studies. It rejected the assumptions by Hunter and Schmidt that supervisory ratings have only a .6 reliability and utilized a .8 reliability for this correction on the basis of the evidence that was available. It also rejected their very generous adjustment of validity scores for the restriction of range in the validation samples caused by preselection of workers according to measures that are correlated with GATB.

Whereas the Committee acknowledged a theoretical need to adjust for restriction of range, in the absence of any plausible empirical data, it decided that no adjustment was the most prudent course of action. The Committee accepted the Hunter and Schmidt (1982) correction for sampling error. The net result of the more recent validation studies and the Committee's corrections is a validity that is less than half of that asserted by the Employment Service.

And (p. 361) ..

The final section of this volume addresses the predicted impact of GATB on economic productivity as well as recommendations on the future use of GATB. One of the major motivations for the Employment Service to use GATB and for employers to place job orders for employees selected by GATB are the economic claims of greater employee productivity using this system of selection. Hunter (1983) claimed gains in economic productivity of about $80 billion in 1983 if GATB were used appropriately for all referrals. He and his colleague, Schmidt, also asserted that the entire economy might gain an additional $153 billion in output if the system were extended to the entire economy (Hunter & Schmidt, 1982). Economists have found these estimates to be wildly exaggerated and based on completely arbitrary and improbable assumptions (Levin, 1989).

The Committee carried out its own analysis and concluded that the probable economic gains would be a mere fraction of those claimed by Hunter and Schmidt (1982). Of particular interest is the fact that Hunter and Schmidt calculated economic gains for the economy by assuming that there are 10 times as many workers applying for jobs as there are available jobs-an assumption of a 90% rate of unemployment. They also calculated economic gains by comparing top-down selection using GATB raw scores with an assumption that workers would be selected at random otherwise. Whereas this will inflate the utility of GATB selection, random selection is almost never the case in employment selection. This assumption is all the more surprising because Hunter and Schmidt inflated substantially the observed validity coefficients for restriction of range based on the assumption that the validation samples had been determined by restrictive selection criteria rather than by random selection

And Levin concludes on page 362:

Finally, it is peculiar that the report did not comment on the implications of the relatively low validity coefficients. According to the findings of this study, only about 5%-6% of the variance in supervisory ratings of workers can be explained by the variance in their GATB scores. This is about the same portion of the variance in workers' earnings that can be explained by test scores. Given the puny explanatory power of test scores for predicting economic outcomes, it is hard to understand why the Committee did not raise the larger question of why alternative approaches to employee selection are not investigated. After finishing this magnificent report, I was still searching for this missing chapter.

This report is yet another powerful reason to rethink our position about the “powerful” predictive utility of cognitive ability tests. Surprisingly, Hunter and Schmidt fail to mention this report in any of their later landmark publications (Schmidt and Hunter, 1998; 2004).

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Predicting Job Performance from GMA - Salgado et al (2004)

Job-Source	Actual mean Validity	"corrected" Validity
Driver	0.22	**0.45**
Electrician	0.28	**0.54**
Information Clerk	0.31	**0.61**
Engineer	0.23	**0.63**
Manager	0.25	**0.67**
Police	0.12	**0.24**
Sales	0.34	**0.66**
Skilled Worker	0.28	**0.55**
Typing	0.23	**0.45**

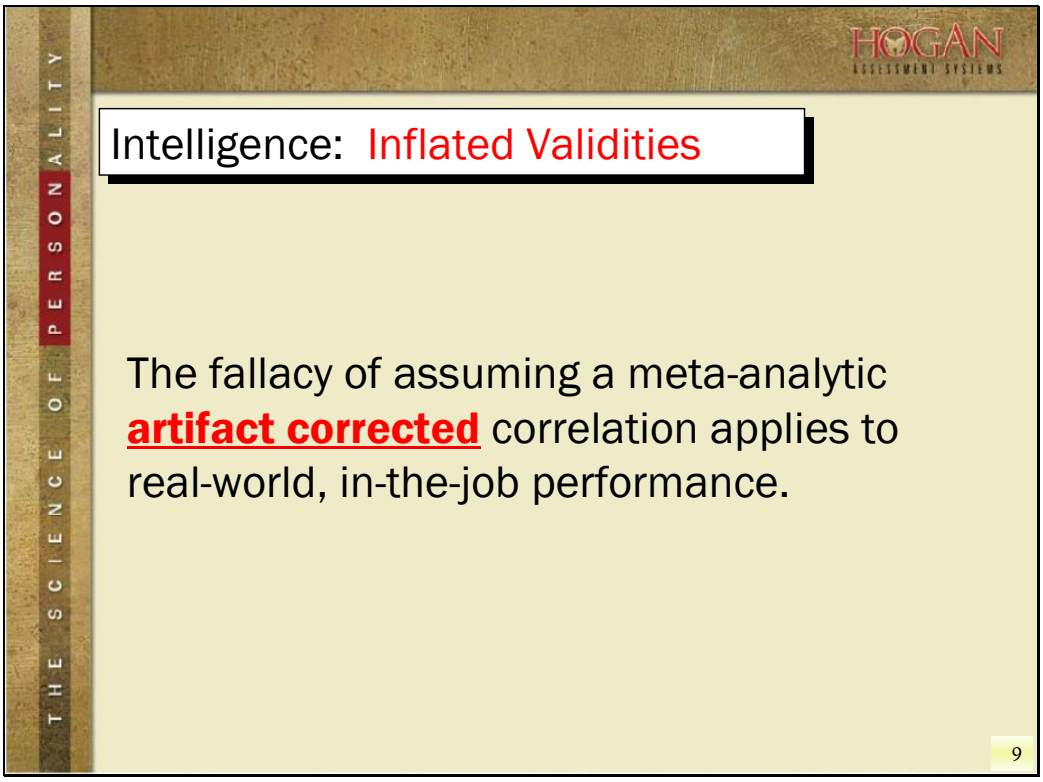
8

In relation to the Hartigan and Wigdor recalculations of validities (as nearly half those reprted by Schmidt and Hunter), it is of interest to look at a table of GMA vs Job Performance ratings from Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J.P. (2003) A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 6, 1068-1081.

We note that the actual mean validity coefficients are nearer to those recalculated by the investigative committee than the boosted, "corrected beyond reality" versions in the second column.

As to those who argue that corrected validities are what matter – the two slides below suggest that this is the case only with regard to initial selection of employees.

It is an unfortunate fact that some I/O psychologists and consultants forget that the relationship assumed to hold for selection still holds for incumbents. It doesn't. This oversight has probably cost more to international industry than all of Schmidt and Hunter's mistakenly assumed utility gains from applying their inflated validities!



The problem, put simply, is that organizations may hire candidates whose ability is above a certain threshold, based upon a validity coefficient which assumes people with IQs less than 30 and above 180 will be applying for the advertised executive position.

Slide #10 shows a data plot of 5000 simulated cases where job performance measured on a 100 point integer rating scale is correlated with GMA (measured using an IQ scale with mean 100 and SD of 15). The bivariate correlation is 0.5, sampling data from a bivariate normal distribution. Note here we're being generous to Schmidt and Hunter, and ignoring the evidence from Hartigan and Wigdor. If we used these validity estimates then the results of the analysis below would be truly depressing!

This value of 0.5 is probably used to convince the buyer of the test that the relationship is pretty good (which it is) - and so selecting new employees using this assessment (GMA) will yield positive benefits. So convinced, an employer might be advised to place a cut-score at say 110 – those scoring above this value being selected for further screening or even an employment offer. Slide #11 on the next page shows the relationship between Job Performance and GMA look like in the now selected job incumbents.

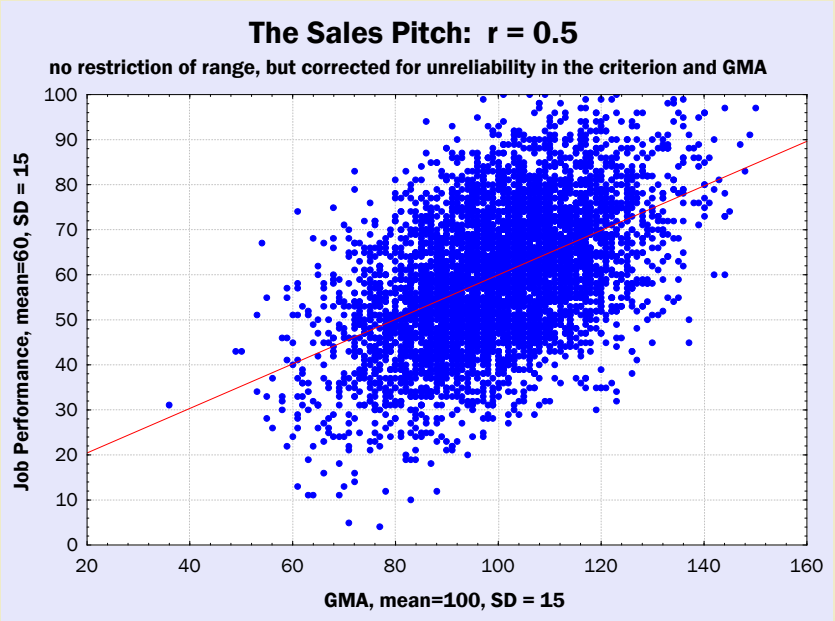
Although it makes sense to use the Validity-Generalized correlation as an indicator of which variables might be important for selection, one must remember that the range-restricted corrections to the observed correlation within incumbents rely upon assumptions of normally distributed data. If these assumptions do not hold, then the estimated correlation will be in error, and the whole exercise rendered flawed. Likewise the corrections for unreliability or predictor and criterion variables.

Our example likely represents the "best-case" scenario, being generated from a theoretical bivariate-normal sampling distribution. There is no such guarantee that real-world data conforms to this distribution.

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Intelligence: Inflated Validities

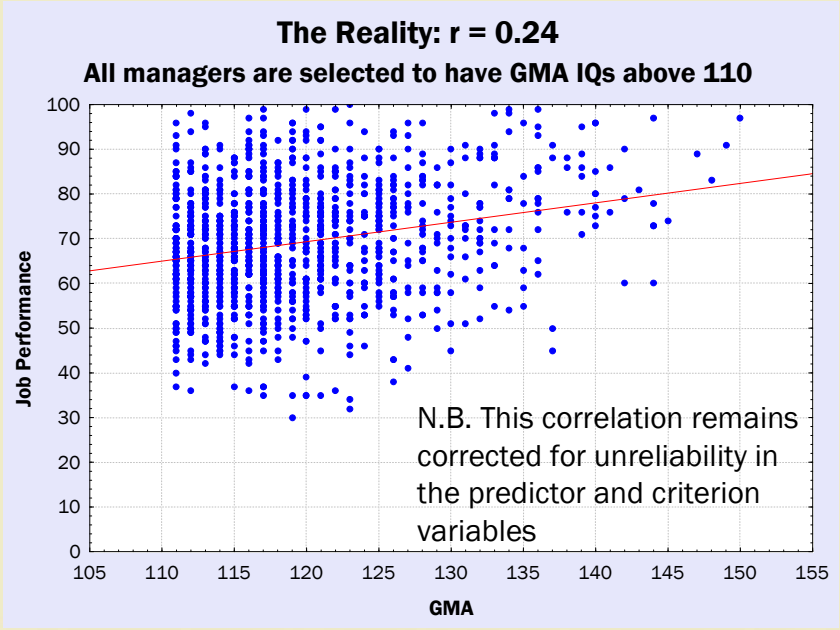


10

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Intelligence: Inflated Validities



11

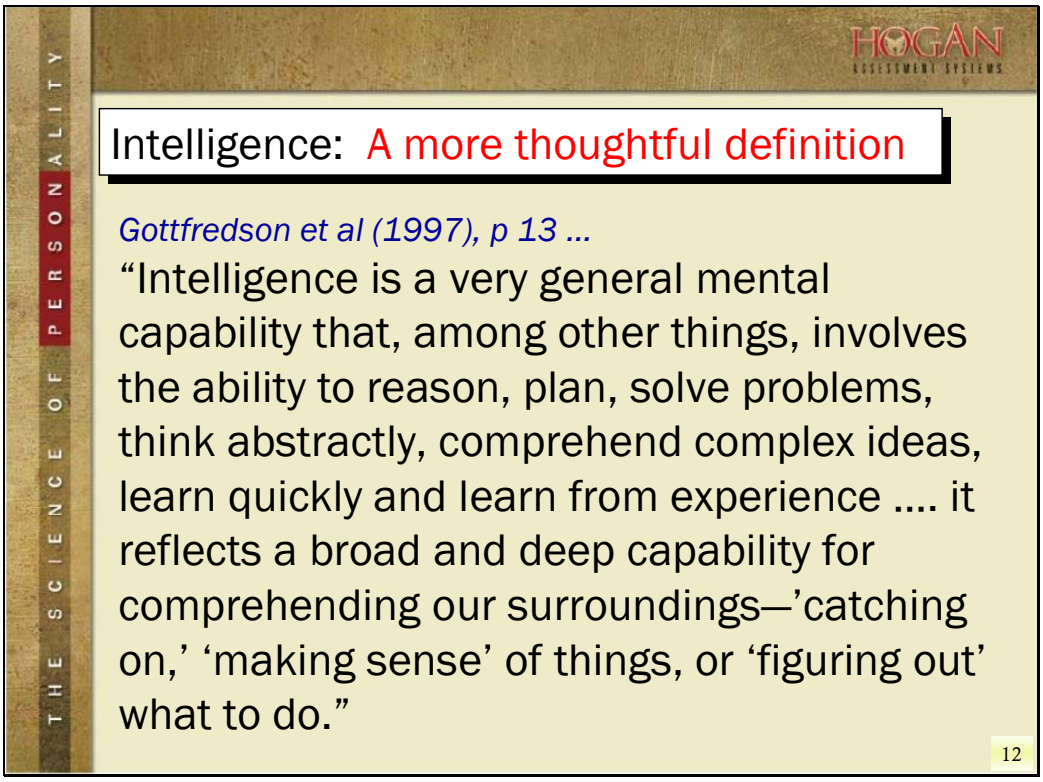
Now you begin to see why variables other than GMA are so important to understanding employee successes and failures – **amongst job incumbents**.

This simple demonstration also reminds us that if we can begin to account for the successes and failures amongst job incumbents with variables beyond GMA, then the selection strategies themselves must change to better reflect those variables important to employee success within an organization.

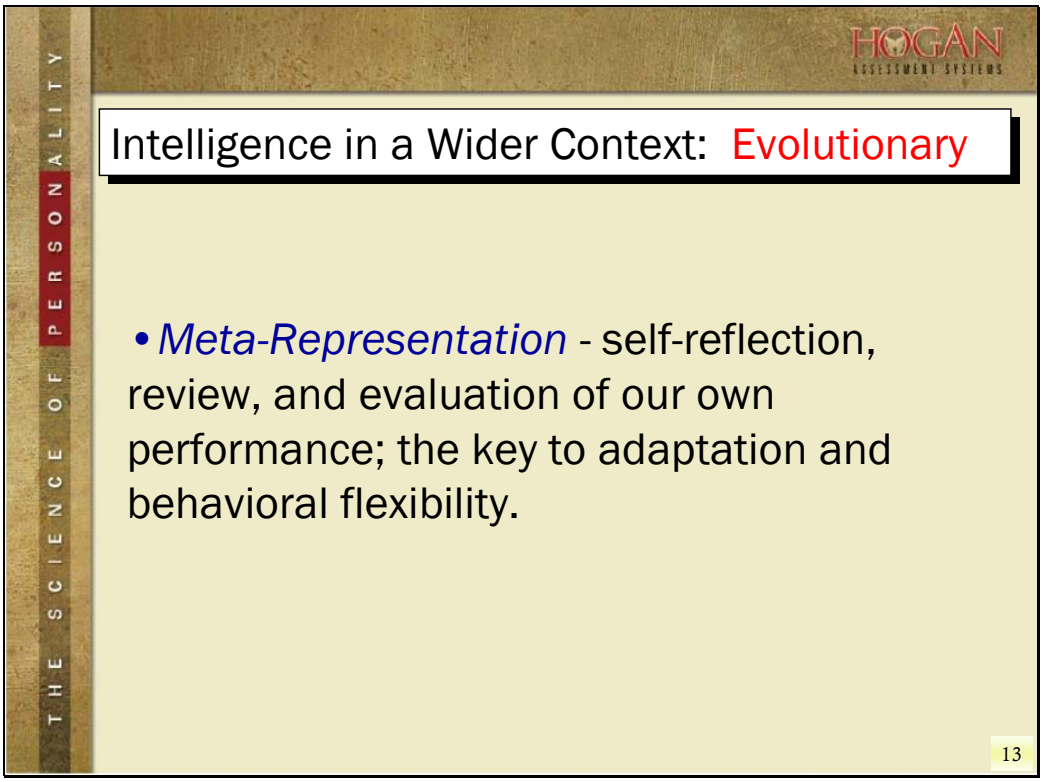
And imagine if the GMA x Job Performance validity is nearer 0.25 … the above incumbent correlation would be near 0.12 (just 1% of the variance in Job Performance accounted for by GMA) – within Job Incumbents).

Let us now briefly examine some definitions of intelligence …

This taken from Gottfredson, L.S. (1997) Mainstream Science on Intelligence: An Editorial With 52 Signatories, History, and Bibliography. *Intelligence*, 24, 1, 13-23.



Then, from our book chapter (Good Judgment: The Intersection of Intelligence and Personality: R.Hogan, J. Hogan, and Barrett in press)



An Evolutionary Model of Intelligence: Meta-Representation

This section outlines our model of intelligence. The word "intelligence" is a recent addition to our language, and it is instructive to note that the ancient Greeks did not use the word. Rather they used words like clever, cunning, and wise to describe individual differences in performance. More importantly, all of these words have behavioral referents—people are only called clever if they routinely manifest a certain kind of performance. In our view, intelligence should be defined in terms of certain behaviors, and people refer to these behaviors when they conclude that someone has "good judgment" or, conversely, "poor judgment."

If the word "intelligence" denotes something real, then it must be rooted in biology and promote individual and group survival—there must be adaptive consequences associated with individual differences in intelligent behavior. In a study of self-consciousness, Sedikides and Skowronski (1997) argue that self-awareness—the ability to think about one's impact on one's social environment—is an adaptive outcome of human evolution. They propose that self-awareness gave early humans an advantage relative to their major competitors.

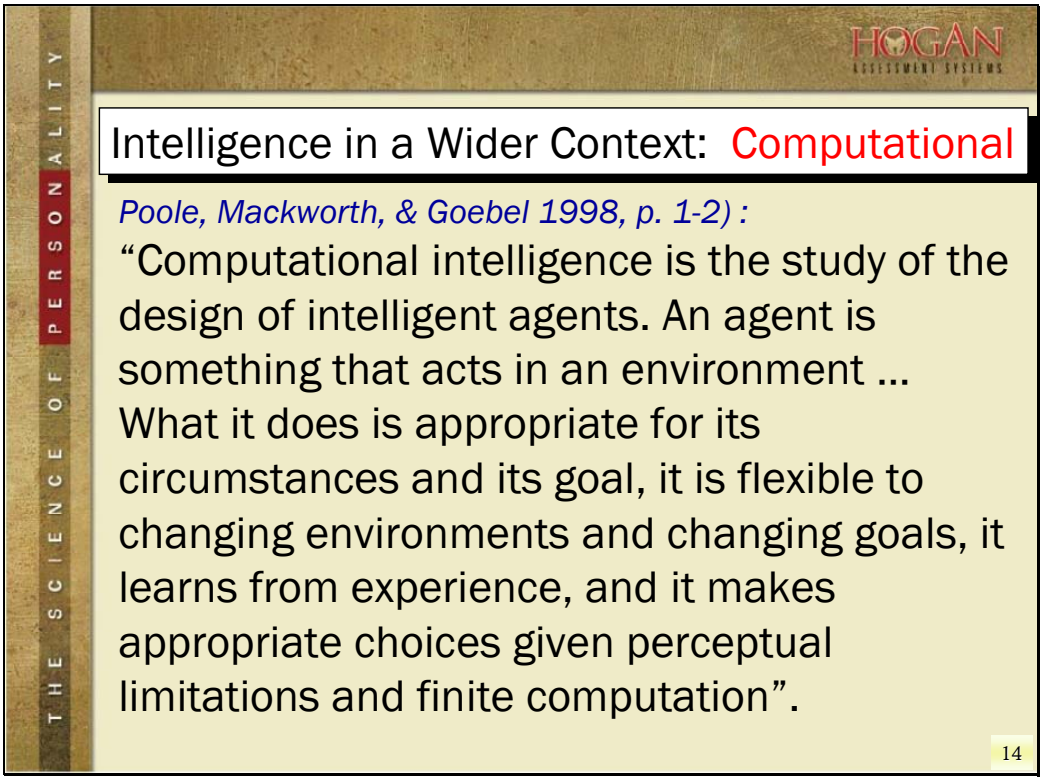
We think that Sedikides and Skowronski (1997) are correct as far as they go—the capacity for self-reflection is a necessary precursor to intelligent behavior. However, we propose that intelligent performance depends on a more general capacity that can be called "meta-representation". By meta-representation, we mean the ability to reflect on our performance (physical, social, intellectual) across all aspects of experience, to review it, and then evaluate it. The definition of stupidity is to continue doing something that yields a poor outcome but to expect that the outcome will improve if one persists in doing the same thing. In contrast, when smart athletes fall behind in a game, they reflect on their performance both on its own terms and

relative to the competition, change their tactics, and then improve their performance—and this is why they are called "smart".

Our hominid ancestors evolved (as group living animals) in an environment that was more demanding and less forgiving than ours. In the ancestral environment, survival depended on being able to solve a wide variety of problems, including finding food, water, shelter, and protection from very nasty predators, keeping the peace within the group, and defending oneself and family against attacks by competing human groups. If the group members did not solve these problems correctly, they died; those that solved the entire range of problems prevailed. But present success is no guarantee of future success. The demands of survival changed constantly; those groups that adapted and improved their survival techniques in the face of constantly shifting environmental pressures became our more recent ancestors—the ultimate winners in the race for survival. Improving one's performance involves correctly anticipating future problems or recalling past performance that yielded better outcomes than those resulting from current performance. In either case, improving performance depends on the capacity for meta-representation, the ability to reflect on and evaluate one's performance, and then use the results of this reflection to improve subsequent performance.

Intelligent people can detect problems in their performance and then change it. They can also detect problems in other peoples' performance and encourage them to change it. Anthropologists and psychologists have traditionally argued that behavioral flexibility is the most important single human characteristic. Meta-representation is the key to behavioral flexibility. Crocodiles are competent hunters and proficient predators. Over time they have eaten many humans, but because crocodilian behavior is largely wired and inflexible, humans can hunt them to extinction.

Then, from the domain of computational intelligence …



From Poole, D., Mackworth, A., and Goebel, R. (1998) Computational Intelligence: A Logical Approach. Oxford University Press. ISBN: 0-19-510270-3, Chapter 1, p. 1-2 ...

"*Computational intelligence is the study of the design of intelligent agents. An agent is something that acts in an environment-it does something. Agents include worms, dogs, thermostats, airplanes, humans, organizations, and society. An intelligent agent is a system that acts intelligently: What it does is appropriate for its circumstances and its goal, it is flexible to changing environments and changing goals, it learns from experience, and it makes appropriate choices given perceptual limitations and finite computation.*

The central scientific goal of computational intelligence is to understand the principles that make intelligent behavior possible, in natural or artificial systems. The main hypothesis is that reasoning is computation. The central engineering goal is to specify methods for the design of useful, intelligent artifacts.

Artificial intelligence (AI) is the established name for the field we have defined as computational intelligence (CI), but the term "artificial intelligence" is a source of much confusion. Is artificial intelligence real intelligence? Perhaps not, just as an artificial pearl is a fake pearl, not a real pearl. "Synthetic intelligence" might be a better name, since, after all, a synthetic pearl may not be a natural pearl but it is a real pearl. However, since we claimed that the central scientific goal is to understand both natural and artificial (or synthetic) systems, we prefer the name "computational intelligence. It also has the advantage of making the computational hypothesis explicit in the name.

The confusion about the field's name can, in part, be attributed to a confounding of the field's purpose with its methodology. The purpose is to understand how intelligent behavior is possible. The methodology is to design, build, and experiment with computational systems that perform tasks commonly viewed as intelligent. Building these artifacts is an essential activity since computational intelligence is, after all, an empirical science; but it shouldn't be confused with the scientific purpose. Another reason for eschewing the adjective "artificial" is that it connotes simulated intelligence. Contrary to another common misunderstanding, the goal is not to simulate intelligence. The goal is to understand real (natural or synthetic) intelligent systems by synthesizing them. A simulation of an earthquake isn't an earthquake; however, we want to actually create intelligence, as you could imagine creating an earthquake. The misunderstanding comes about because most simulations are now carried out on computers. However, you shall see that the digital computer, the archetype of an interpreted automatic, formal, symbol-manipulation system, is a tool unlike any other: It can produce the real thing.

The obvious intelligent agent is the human being. Many of us feel that dogs are intelligent, but we wouldn't say that worms, insects, or bacteria are intelligent (Exercise 1.1). There is a class of intelligent agents that may be more intelligent than humans, and that is the class of organizations. Ant colonies are the prototypical example of organizations. Each individual ant may not be very intelligent, but an ant colony can act more intelligently than any individual ant. The colony can discover food and exploit it very effectively as well as adapt to changing circumstances. Similarly, companies can develop, manufacture, and distribute products where the sum of the skills required is much more than any individual could understand. Modem computers, from the low level hardware to high-level software, are more complicated than can be understood by any human, yet they are manufactured daily by organizations of humans. Human society viewed as an agent is probably the most intelligent agent known. We take inspiration from both biological and organizational examples of intelligence "

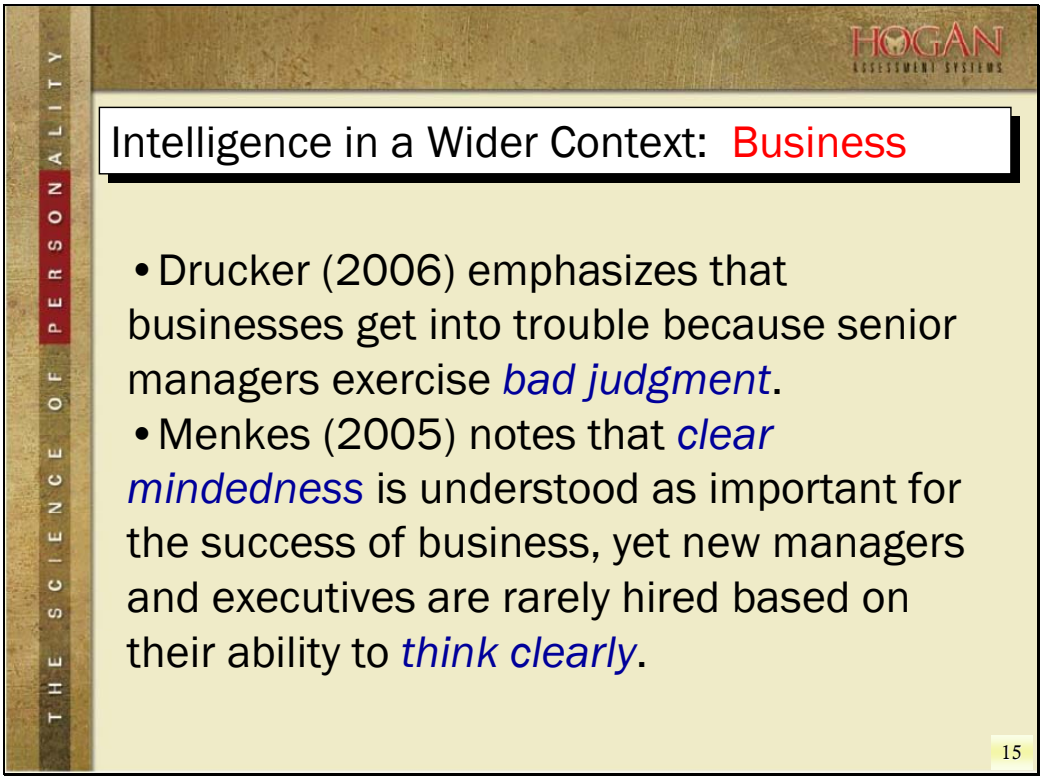
Therein lies the distinction between a view of mental abilities and indeed IQ (as mental "componentry") as "indicators of differential human performance" – vs what might be called the Hogan and Barrett view of intelligence – which is virtually synonymous with the entirely abstract definition from Poole et al given the above "*An intelligent agent …: What it does is appropriate for its circumstances and its goal, it is flexible to changing environments and changing goals, it learns from experience, and it makes appropriate choices given perceptual limitations and finite computation.*"

So, our rejoinder to those who would claim that "intelligence" is a "make it up as you go" and fuzzy concept is that contrary to this view, the field of computational intelligence has a very clear definition of intelligence which accords directly with psychology's own evolutionary theory of the function of "intelligence" within an organism.

What has happened is that the mental component/factor-analytic model view of mental abilities has gradually become less and less relevant to anything other than the statistical taxometric modeling of, specific mental abilities. If that is all we wish to address, then Jensen, Hunt, and Carlson are absolutely correct to state that the word "intelligence" has no place in this exercise.

But. within employment settings, unless we are employing individuals who are required to possess high levels of very specific skills or mental abilities (because they need these high levels of skills/abilities to complete the work required of them), then we are indeed in the realm of the Hogan and Barrett model – where looking for "Intelligent" individuals with "Good Judgment" becomes paramount.

Finally, it is instructive to look at how others from the world of business and organizational philosophy view the issues …



Peter Drucker, the fabled philosopher of management, constantly emphasized that businesses get in trouble because senior managers exercise bad judgment (cf. Drucker, 2006). Managers are supposed to

direct money and energy toward activities that increase profitability. More often, however, they spend time and money solving problems and completing projects that don't matter. It takes clear minded analysis to determine how appropriately to use money and energy. Clear mindedness is a function of good judgment.

Drucker also emphasized that virtually every major business crisis results from the fact that the assumptions on which the business was built and is being run no longer fit reality. Drucker called these assumptions the "theory of the business". Constructing a valid theory of the business, and then subsequently evaluating and revising it, is a function of good judgment.

Menkes (2005) notes that, although most people understand the importance of clear mindedness for the success of business, new managers and executives are rarely hired based on their ability to think clearly. There is a clear need for sound and defensible methods to evaluate the ability of managers and executives to think clearly and exercise good judgment.

Looking back at the history of intelligence theorizing and research, Hogan et al (in press) noted that a two-component model seemed to be spoken of, or capable of being derived, from many different individual's work in the area. The next slide shows some of these contributions:

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Source	Dimension 1	Dimension 2
R. Hogan	**Problem Finding**	**Problem Solving**
C. Spearman	Eduction	Reproduction
	(Problem Finding/Solving)	(Describing the Solution)
C. S. Peirce	Forming Hypotheses	Evaluating Hypotheses
L. L. Thurstone	Rule Finding	Rule Applying
R. B. Cattell	Fluid Intelligence	Crystallized Intelligence
J. P. Guilford	Divergent Thinking	Memory/Cognition/ Convergent Thinking
	Evaluation	
J. Piaget	Formal Operations	Concrete Operations
H. Reichenback	Context of Discovery	Context of Justification
B. Haig	Phenomena Detection	Theory Construction
Business Speak	**Strategic Reasoning**	**Tactical Reasoning**

16

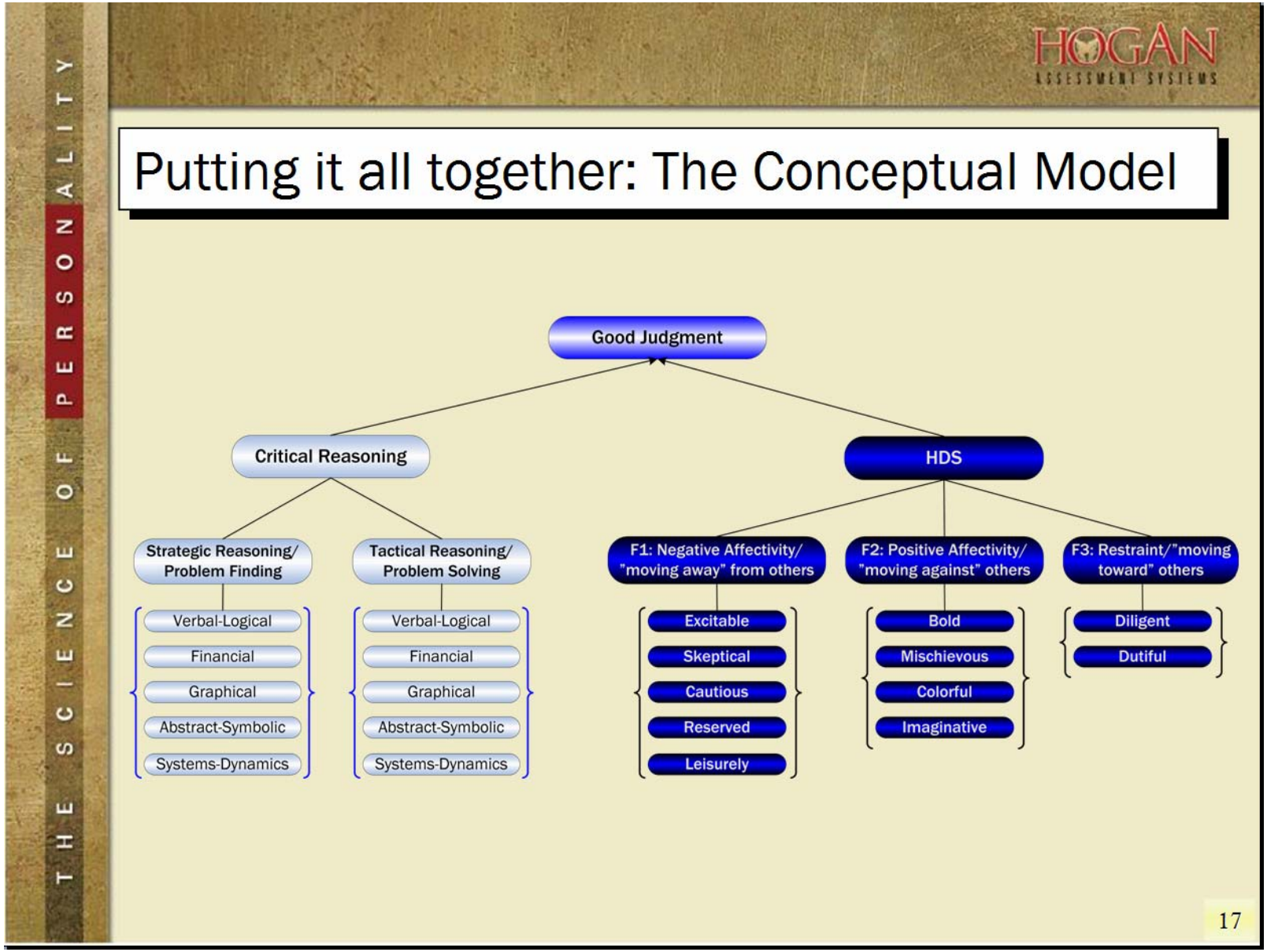
It is also clear that we have drawn upon some of the thinking from Newell, Shaw and Simon (1958) and Newell and Simon (1970) in problem solving, and the notion of problem detection first mooted by Smith (1989) and leter expanded upon by Klein, Pliske, Crandall, and Woods et al (2005).

From Hogan et al (in press):

> Standardized intelligence testing was developed to predict academic performance (Binet & Simon, 1905).
> In contrast, an effective measure of executive intelligence should predict clear thinking, good judgment, and effective management decision making. Drucker describes the key components of this in very general terms: thinking critically about the theory of the business by reviewing the assumptions on which it was founded and in terms of which it is being operated. We believe this process can be usefully specified in terms of time perspective as follows:
>
> 1. Past perspective: Are the operating assumptions of the business still valid?
> 2. Present perspective: Given the stated goals of the business, are people currently working on the right problems and tasks?
> 3. Future perspective: Given the stated goal of the business, have people appropriately anticipated the potential future problems and possible future outcomes correctly?
>
> Within each of these perspectives two kinds of thinking will apply.
> We call them "**problem-finding**" and "**problem-solving**"; we also refer to these two kinds of thinking as "**strategic reasoning**" and "**tactical reasoning**".
>
> - Problem-finding involves detecting gaps, errors, or inconsistencies in data, trends, textual materials, existing processes and procedures, or verbal arguments.
>
> - Problem solving involves finding answers to the problems once they are identified, following arguments to their logical conclusions, and applying well-understood methods to new problem categories.
>
> - Our view of competent business reasoning and good judgment starts with the preceding discussion. It assumes that the word "intelligence" refers to clear thinking and is a key component of successful managerial performance.
>
> - It assumes that intelligence facilitates good judgment.
>
> - It assumes that two kinds of reasoning are essential to this process—problem finding and problem solving.
>
> - It assumes that these two kinds of thinking can be measured, and that the results from this measurement process can be used to evaluate good judgment.
>
> - It assumes that good judgment is a function of both intelligence and personality, that personality can be measured, and that a person's reasoning style can be best estimated using measures of intelligence and personality.
>
> - And finally, it assumes that the results of this measurement process will predict successful occupational performance.

The model which is driving the development of the latest version of the Hogan Business Reasoning Inventory (HBRI) is given below. The Hogan Development Survey (HDS) is the measure of personality attributes which we feel is most relevant to the theory of "Good Judgment" and "clear thinking".

And finally …


THE SCIENCE OF PERSONALITY
HOGAN
ASSESSMENT SYSTEMS
Putting it all together: The Conceptual Model
Good Judgment
Critical Reasoning
HDS
Strategic Reasoning/ Problem Finding
Tactical Reasoning/ Problem Solving
F1: Negative Affectivity/ "moving away" from others
F2: Positive Affectivity/ "moving against" others
F3: Restraint/"moving toward" others
Verbal-Logical
Financial
Graphical
Abstract-Symbolic
Systems-Dynamics
Verbal-Logical
Financial
Graphical
Abstract-Symbolic
Systems-Dynamics
Excitable
Skeptical
Cautious
Reserved
Leisurely
Bold
Mischievous
Colorful
Imaginative
Diligent
Dutiful
17

References

Browne-Miller, A. (1995). *Intelligence Policy: Its impact on college admissions and other social policies*. New York: Plenum Press.

Drucker, P. F. (2006). What executives should remember. *Harvard Business Review*, 84, 145-152.

Fuillet, L., Dufour, H., & Pelletier, G. (2007) Brain of a white-collar worker. *The Lancet*, 370, 3, 262.

Gottfredson, L.S. et al (1997) Mainstream Science on Intelligence: An Editorial With 52 Signatories, History, and Bibliography. *Intelligence*, 24, 1, 13-23.

Hartigan, J.A., & Wigdor, A.K. (1989) *Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery.* National Academy Press. ISBN: .

Hogan, R., Hogan, J., & Barrett, P. (in press) Good Judgment: The Intersection of Intelligence and Personality. In G. Bedney (Editor), *Ergonomics and Psychology: Developments in Theory and Practice*. London: Taylor & Francis.

Hunt, E., Carlson, J. (2007a) Considerations relating to the study of group differences in intelligence. *Perspectives on Psychological Science*, 2, 2, 194-213.

Hunt, E., Carlson, J. (2007b) Reply to commentators. *Perspectives on Psychological Science*, 2, 2, 224-226.

Jensen, A.R. (1998) *The g Factor: The Science of Mental Ability*. Praeger,

Klein, G., Pliske, R., Crandall, B., & Woods, D.D. (2005) Problem Detection. Cognition, *Technology and Work*, 7, 14-28.

Levin, H.M. (1991) Fairness in Employment Testing: Validity Generalization, Minority Issues, and the General Aptitude Test Battery - A review. *Journal of Educational Measurement*, 28, 4, 358-363.

Menkes, J. (2005). Executive intelligence. New York: HarperCollins.

Newell, A., Shaw, J.C., & Simon, H.A. (1958) Elements of a theory of human problem solving. P*sychological Review*, 65, 3, 151-166.

Poole, D., Mackworth, A., & Goebel, R. (1998) *Computational Intelligence: A Logical Approach*. Oxford University Press.

Richters, J.E. (1997) The Hubble hypothesis and the developmentalist's dilemma. *Development and Psychopathology*, 1, 2, 193-229.

Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rolland, J.P. (2003) A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 6, 1068-1081.

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. (2004) General mental ability in the world of work: occupational attainment and Job Performance. *Journal of Personality and Social Psychology*, 88, 6, 162-173.

Simon, H.A., Newell, A. (1971) Human problem solving: the state of the theory in 1970. *American Psychologist*, 26, 2, 145-159.

Smith GF (1989) Defining managerial problems: a framework for prescriptive theorizing. *Management Science* 35(8):963–981.

Sternberg, R.J., Grigorenko, E. (2007) The difficulty of escaping preconceptions in writing an article about the difficulty of escaping preconceptions: commentary on Hunt and Carlson (2007). *Perspectives on Psychological Science*, 2, 2, 221-223.

Stankov, L. (2003) Complexity in Human Intelligence. In R.J. Sternberg, Lautrey, J., and Lubart, T.I. (eds) *Models of Intelligence: International Perspectives* (Chapter 2, pp 27-42). Washington DC: American Psychological Association.

The Economist, Science and technology Section (2007) *Really New Advances*. June 14th, pp 87-89.