THE SCIENCE OF PERSONALITY

Two Big Ideas

And their Consequences

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

Hogan Assessment Systems Inc.
Tulsa, Oklahoma, USA

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

#1: Traits

From: Hogan, R. (2005) In defense of personality measurement: new wine for old whiners. *Human Performance*, 18, 4, 331-341 … pp. 334-335

3

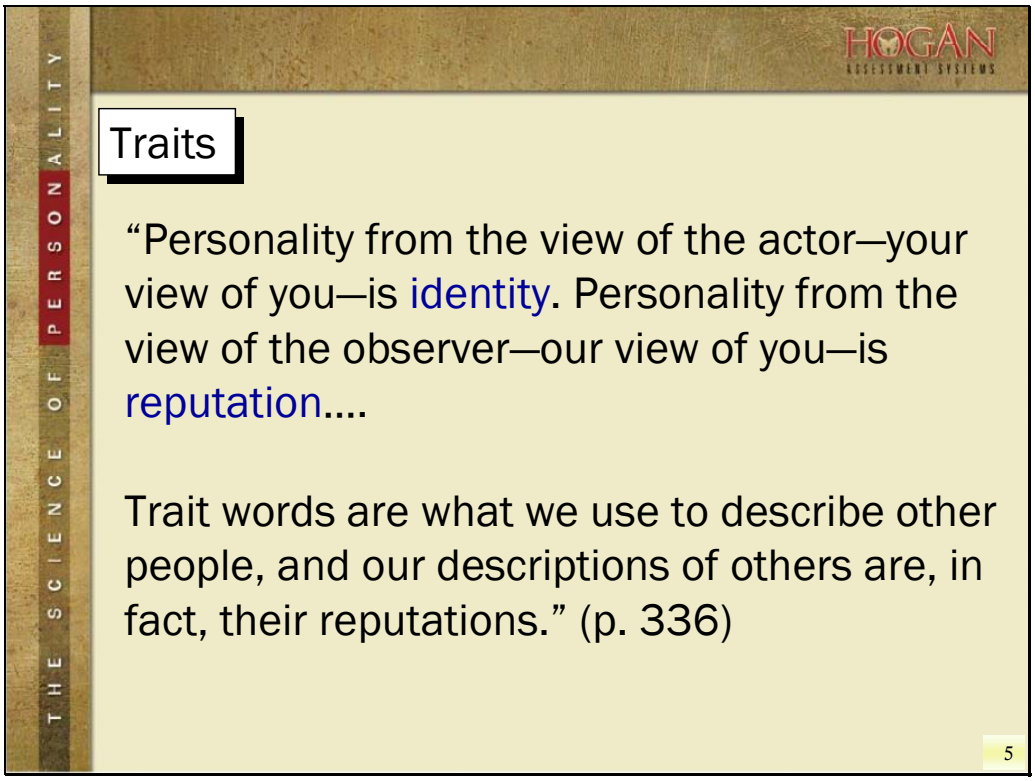
THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

Traits

"However, other people don't have traits; rather, we assign trait terms to them as a way of summarizing recurring themes in their behavior. There is a difference between description and explanation, and trait theorists ignore the distinction".

4



When I (Paul Barrett) first read this last year, I nearly fell off my seat. I had spent the past 25 years or so embedded firmly within a "traits as causal entities" – with biological causation implied as a kind of biophysical substrate of the observed trait behaviors. In short, a typical Big 5 and Eysenckian view of personality. Whilst there is undeniably a biological cause to the way we approach stimuli and respond to them "stylistically", I think evolutionary psychology has rendered this view of traits as "fundamental" psychological "entities" as impossibly simplistic.

My view of "traits" as "neuropsychic entities" (as Bob puts it) was already in question after reading Michael Maraun's perceptions and arguments contained with a mini-debate with Marvin Zuckerman on the nature of sensation–seeking and traits per se, as well as his demonstration that the Big Five Model was not a fundamental psychological structure at all, but merely what you get from a linear covariance decomposition. Change the metric and method of analysis, and you produce another structure altogether.

I've included the abstracts to each paper here – as I cannot stress how important each is in its own right.

Jackson, J.S.H., & Maraun, M. (1996) The conceptual validity of empirical scale construction: the case of the Sensation Seeking Scale. *Personality and Individual Differences*, 21, 1, 103-110.

> *An assessment of the present state of sensation seeking (SS) research suggests that the concepts of physical risk taking and sensation seeking have become synonymous with version five of Zuckerman's Sensation Seeking Scale (SSS). Given this, it would appear that the validity of SS research rests on the extent to which the SSS actually measures risk taking and sensation seeking tendencies. The present paper suggests that a series of logical and conceptual errors in the development of the SSS have resulted in a scale that does not measure sensation seeking. This raises concerns about both the value and meaning of investigations in which it is employed.*

It is concluded that risk taking and SS research has learned more about the SSS than sensation seeking itself.

Zuckerman, M. (1996) "Conceptual Clarification" or confusion in "The study of Sensation Seeking" by Jackson and Maraun. *Personality and Individual Differences*, 21, 1, 111-114.

Jackson and Maraun claim that the Sensation Seeking Scales (forms II, IV, and V) do not measure sensation seeking because the scales were developed using factor analyses of items rather than rationally using the original construct. The constructs were changed after the factor analyses revealed a different structure than anticipated and the new constructs were validated using Cronbach and Meehl's 'construct validity'. One cannot assume that the meanings attributed to test items on a rational basis is the same as that given by test respondents. If there are subfactors one cannot assume they are organized in the meaning dimensions hypothesized by the test constructors. Meaning resides in the construct validity of the test as well as the content validity and factor analyses are an important step in establishing content validity.

Jackson, J.S.H., & Maraun, M. (1996) Whereof one cannot speak, thereof one must remain silent. *Personality and Individual Differences*, 21, 1, 115-118.

When an individual claims that "these are measurements of σ, or "the way to measure the σ of δ is. .. ", his claims are not made correct or incorrect via the products of empirical investigation. Regardless of whether σ is a physical or psychological concept, the justification for such claims comes from: (1) the existence of a rule guided practice for measuring σ; and (2) having correctly followed the rules that govern the practice. Rules (standards of correctness) are a precondition to distinguishing between claims that are correct and claims that are incorrect. Why is this the case? A measurement claim has the form: the σ of δ is x, in which x is standardly given in particular units. The practice of measuring the σ of δ in particular units is the basis for a grammatical relation between σ and δ, for it rests on the meanings of both σ and δ, and meaning is manifest in the rules for the correct use of concepts (Wittgenstein, 1953; Ter Hark, 1990). The justification for the statement that voltmeters measure voltage, not electricity, is given in comparison to rules that establish the meaning of electricity and voltage. Similarly, the rejection of the sensation seeking scale (SSS) as a measure of sensation seeking follows from careful scrutiny of the grounds for the instantiation of a disposition, in contrast to a need, preference, attitude or desire (see for example: Alston, 1975; Wittgenstein, 1953; Ter Hark, 1999; Rozeboom, 1984; Tuomela, 1978). Let us briefly review relevant theory.

The paper closes with:

The problems of psychology are some of the most difficult and worthwhile in science. But a much greater sophistication is needed if we are to have the productive, respected science we all desire. As Norris (1983, p. 53) states in discussing construct validation theory, "there is a lack of penetrating criticism and of alternative views in discussions of the field." This dogmatic allegiance to Cronbach and Meehl's archaic treatment of measurement will not do. We must make getting it right, rather than getting it published, the guiding principle of psychological investigation.

And finally ...

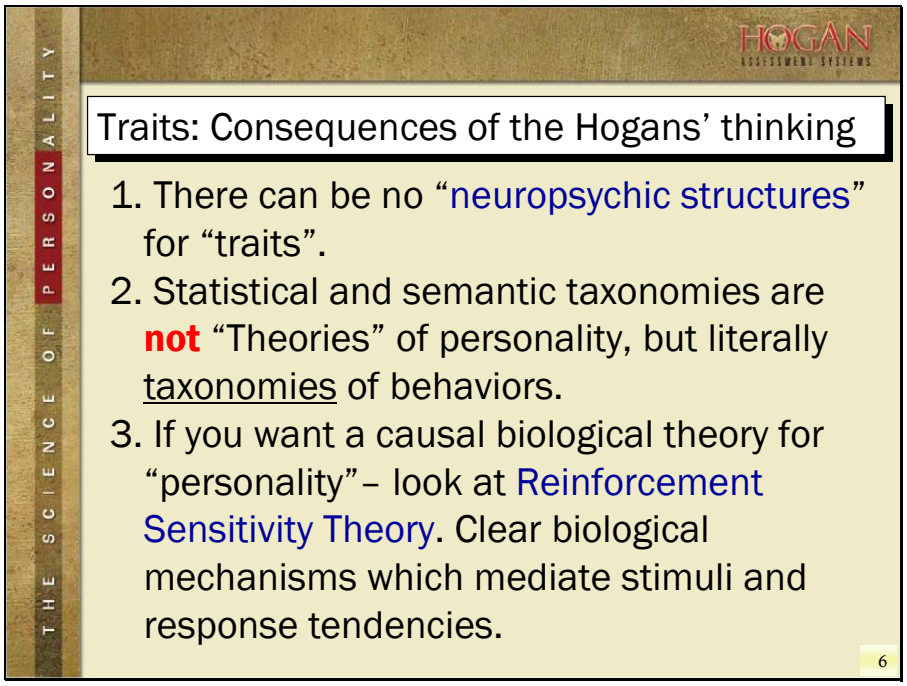
Maraun, M.D. (1997) Appearance and Reality: Is the Big Five the Structure of Trait Descriptors? *Personality and Individual Differences*, 22, 5, 629-647.

It is argued that contrary to the claims of Big Five investigators, the structure of trait descriptors is still very much an open issue. This is because their methodology, factor/component analysis

paired with the dimensional interpretation/simple structure (DISS) procedure, does not investigate the closed topological manifold that constitutes the 'structure' of a set of variables. Instead, radex-related configurations are likely candidates for the structure of trait descriptors. Some preliminary support for this claim is given by an analysis of the NEO Personality Inventory (NEO-PI) (Costa & McCrae, Manual for fhe NEO PI-R. Odessa, FL: Psychological Assessment Resources, 1992), a Big Five questionnaire measure, and the Goldberg-40 (Goldberg, Review of Personality & and Social Psychology, Vol. 2, pp. 141-165. 1981. Beverly Hills, CA: Sage), an adjective measure. In particular, the NEO-PI and Goldberg-40 are shown to have radex structures. A facet theory (Guttman, Psychometrika, 36, 329-346, 1971) rationale is provided for these findings.

But, I digress. Bob's paper and Joyce Hogan's writings in this area suddenly made a great deal of sense. During the late 80s and 90s, I had also become aware of Jeffrey Gray's Reinforcement Sensitivity Theory of Personality. There are two excellent reviews of the newer incarnation of this theory – one by Smillie, Pickering, & Jackson (2006) *The new Reinforcement Sensitivity Theory: implications for personality measurement*, and one by Bill Revelle (in press). *The Contribution of Reinforcement Sensitivity Theory to Personality Theory*. The most important aspect of RST is that it explains behavior in terms of three fundamental but interacting biological systems – these are not trait-systems – but broad physiological systems which respond to environmental stimuli, generating motivation, internal cognitions, and behaviors. Disentangling these systems is a real headache – but, as noted below under Big Idea #2, this may be impossible.

Although no author mentioned above probably feels this way, I think Bob and Joyce's socioanalytic theory really lays bare that traits are not "fundamental biological entities at all" – which is why the RST field is stuck trying to figure out how to bolt psychometric conceptualizations of traits onto RST dynamic stimulus-response/learning systems. Rather, trait systems are simply taxonomies of behavior – not theories of behavior and not explanatory concepts.



If you take the big idea from the Hogans … then "trait psychology" disappears as really all we are working with are self-reports of reputations, ordered into convenient categories which have some valuable predictive value. The behaviors repored as "being like them", are the result of a mixture of biological influences and cognitions. We simply categorize these into a reasonable taxonomy. We might even say people vary in terms of the magnitude of reputation variable they wish to make known to others. But, we are a long way from assuming these taxonomic variables possess a quantitative structure.

Which is why simple sum scores of magnitudes from self-report items apparently predict just as much behavioral variance as do IRT, and Ideal Point mathematically-sophisticated psychometrics (Cherneyshenko, Stark, Drasgow, F. & Roberts, B.W. (2007) *Constructing personality scales under the assumption of an Ideal Point response process: toward increasing the flexibility of personality measures*. The key to prediction improvements is unlikely to be found in sophisticated mathematics – but in recognizing the very broad nature of the concepts we are using (descriptors of loosely associated behaviors which are subsumed under "trait" rubrics such as Adjustment, Prudence, Extraversion, Conscientiousness). A point made clearly by Maraun (1998), and earlier by Schonemann, (1994) *Measurement: the Reasonable Ineffectiveness of Mathematics in the Social Sciences*. This last paper is a "must-read" for any researcher thinking that the application of advanced mathematical methods is what's required to "do better". Basically this argument is dead in the water before it even begins.

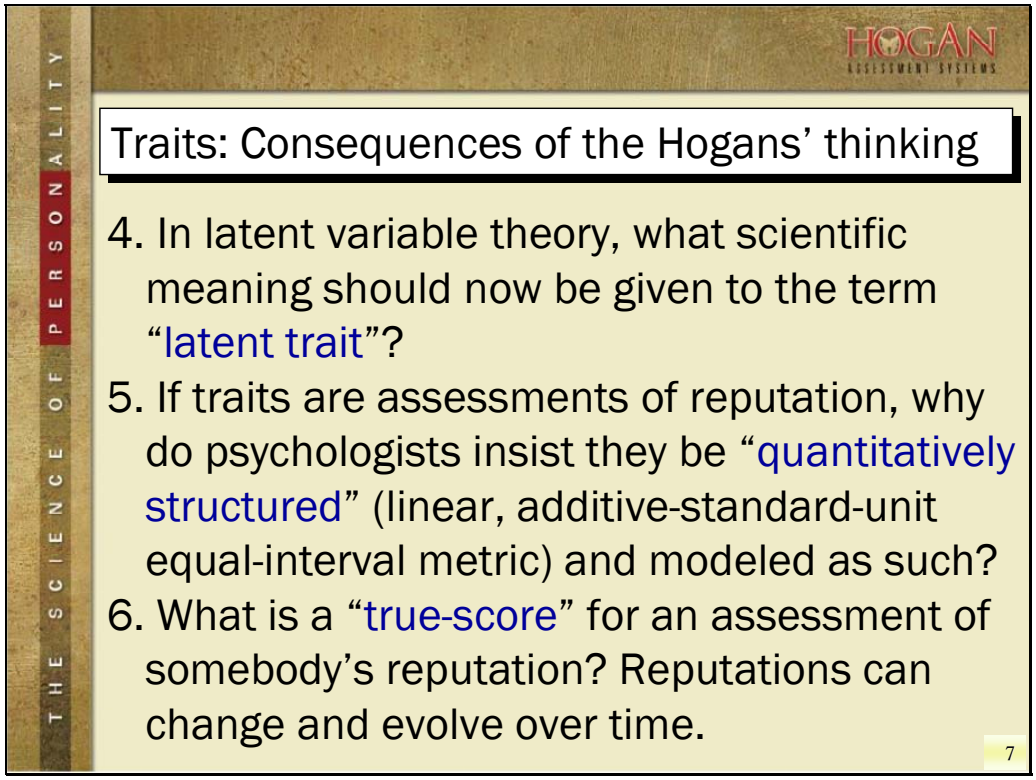
The system we are observing (us, our brains) is not a piece of physical push-pull hydraulics or simple systems causality. That is, the system (us) is both subject to external causal influences as well as an internal originator of cause itself. This is quite unlike the world of physics. It's akin to a billiard ball being struck, then deciding for itself halfway through its trajectory that it doesn't want to go in that direction, and thus changing its response! What's more, this neural system, as noted in the Big Idea #2 below, seems to be adaptive and self-organizing. It changes over time in response to its environment. Note well that recent Lancet article we reported upon in our presentation entitled "Good Judgment" …Fuillet, L., Dufour, H., & Pelletier, G. (2007) Brain of a white-collar worker. *The Lancet*, 370, 3, 262.

This plasticity is also a feature of the causal model proposed by Garlick (2002) for many findings in intelligence research ..

> *The nature of the general factor of intelligence, or g, is examined. This article begins by observing that the finding of a general factor of intelligence appears to be inconsistent with current findings in neuroscience and cognitive science, where specific connections are argued to be critical for different intellectual abilities and the brain is argued to develop these connections in response to environmental stimuli. However, it is then observed that if people differed in neural plasticity, or the ability to adapt their connections to the environment, then those highly developed in one intellectual ability would be highly developed in other intellectual abilities as well. Simulations are then used to confirm that such a pattern would be obtained. Such a model is also shown to account for many other findings in the field of intelligence that are currently unexplained. A critical period for intellectual development is then emphasized.*

So, traits as "neuropsychic entitites" simply don't pass muster anymore.

Which begs the following questions …



The entire basis for much of psychometric theory – both classical and modern – is now frankly questionable. It's one thing to suggest that maybe variables can be constructed under a set of convenient measurement assumptions, then seek to test whether these assumptions hold in reality via experimental observation. But, it's quite another to maintain the construction effort without ever considering whether the system under examination could ever produce behavior or such regularity and linear precision.

If traits are merely variables which allow us to assess people's reputations, then why should be expect equal-interval precision at all? Look at the abstract of a forthcoming paper from Guenter Trendler (Theory and Psychology) … *Measurement Theory, Psychology and the Revolution That Cannot Happen*

> *Doubt is raised that revolutions in measurement theory, for example conjoint measurement, will lead to the quantification of psychological attributes. First the meaning of measurement is explained. Relying on this it is argued that in order to quantify intensive quantities we must firstly, manipulate the attributes involved and, secondly, we must control potential systematic disturbances. It is emphasized that given the causal complexity of the world we must construct experimental apparatus to accomplish these tasks. The creation of modern quantitative science through the adoption of this method is called the Galilean revolution. Next the Millean quantity objection is formulated: it is maintained that if the Galilean revolution is not possible in psychology, then psychological attributes are not measurable. The objection is defended. It is shown that psychological attributes are indeed neither manipulable nor controllable to the required extent. For these reasons it is concluded that they are not measurable.*

In this context, Denny Borsboom's (1966) paper "The Attack of the Psychometricians" looks more like the last trumpet call-to-arms of those engaged in a hopeless cause than a reasoned set of arguments from

scientists carefully exploring phenomena. Especially since Joel Michel's critique of IRT theory still stands unanswered. From: Michell (2004) …p. 126 …

> *Probabilistic item response models are usually expressed in terms of the probability of a person making a response of a certain kind to a test item, given the relevant person and item parameters. Considering one-parameter ability test models, Sutcliffe (1986, p. 91) suggests thinking of a person's correct or incorrect item performance on any occasion as due to a person parameter he calls capability, which is ability, as normally understood in these models, plus a random error component, notionally drawn on each occasion from a probability distribution of possible 'errors', normal, for example, in normal ogive item response models (Lord, 1952) and of an inverse hyperbolic tangent form for logistic ogive models (Rasch, 1960). When item difficulty exceeds capability, the response is incorrect, otherwise it is correct. Now, if a person's correct response to an item depended solely on ability, with no random 'error' component involved, one would only learn the ordinal fact that that person's ability at least matches the difficulty level of the item. Item response modellers derive all quantitative information (as distinct from merely ordinal) from the distributional properties of the random 'error' component. If the model is true, the shape of the 'error' distribution reflects the quantitative structure of the attribute, but if the attribute is not quantitative, the supposed shape of 'error' only projects the image of a fictitious quantitivity. Here, as elsewhere, psychometricians derive what they want most (measures) from what they know least (the shape of 'error') by presuming to already know it.*
>
> *If the random 'error' concept is retained, but it is admitted that the shape of these 'errors' is unknown, then at best only ordinal relationships between people (or items) follow from test performances (Grayson, 1988) unless the cancellation conditions alluded to above (namely double cancellation, triple cancellation, etc.) obtain.* ***By any fair-minded reckoning, our state of knowledge at present does not support claims to be able to measure psychological attributes using item response models.***

However, let's look at the implications of the Hogans' thinking on traits from another completely different perspective; that of the physics of complex systems.

Before we go to big idea #2 – I want to "reprint" a section from Bob's 2005 paper, "In defense of personality measurement: new wine for old whiners", as this is perhaps one of the clearest expositions of what's wrong with "trait" psychology, and where we find ourselves when wishing to deal with clearly observable regularities in behaviors among individuals… (pp. 334-336)

WHAT IS PERSONALITY?

> *Everyone has a theory of personality; we can't go to work without one. The problem is that these theories are informal, implicit, and unspecified. The same is true for industrial psychologists; although they use personality measures, their theory of personality is rudimentary. That is, they define personality as traits, and that is a mistake because trait theory has significant flaws. The problems can be quickly outlined. On the one hand, it confuses description with explanation, and is, therefore, completely tautological. For example, Mike Tyson is usually described as aggressive. Trait theorists want to explain Tyson's aggressive behavior in terms of a trait for aggression, and that's just dumb. Sophisticated trait theorists try to escape from the tautology by arguing that regularities in behavior are caused (and explained) by underlying "neuropsychic structures." This is psychological reductionism, an effort to explain phenomena at one level in terms of phenomena at the next lower level of analysis. Thus, biology should be reduced to the laws of chemistry, chemistry should be reduced to the laws of nuclear physics, and overt behavior should be reduced to neuropsychic structures.*
>
> *There are two problems with reductionism. On the one hand, the rest of science has simply moved on. For example, the Nobel Prize–winning physicist P. W. Anderson (1972) argued 30 years ago that biologists can study biology without worrying about chemistry, that chemists can*

study chemistry without worrying about physics, and so on. More recently, string theory—the theory of everything—is intended to make Einstein's relativity theory (it's about gravity) consistent with quantum mechanics (particles that have no gravitational properties). But some physicists think that relativity and quantum mechanics are separate disciplines that can be studied fruitfully on their own. Similarly, we can study occupational performance without resorting to physiology.

The second problem with reductionism in personality psychology is that, after 70 years, we still haven't found any underlying neuropsychic structures—and don't hold your breath. Obviously people are biological animals, and our actions reflect our genetic makeup, but that is all we need to say. Neuroscientists can study neuropsychology, and applied psychologists can study social behavior on its own terms. The bottom line is that trait theory is not a competent theory of personality. It is a mistake to confuse the way we use trait words with trait theory. Trait words are indispensable for describing other people. However, other people don't have traits; rather, we assign trait terms to them as away of summarizing recurring themes in their behavior. There is a difference between description and explanation, and trait theorists ignore the distinction. We describe other peoples' behavior with trait words, but we explain their behavior in terms of what they are trying to accomplish.

Personality is two things: (a) generalizations about human nature, and (b) explorations of individual differences. What generalizations can we make about human nature? Sociology, anthropology, and evolutionary psychology suggest three. First, people always live in groups. Second, every group has a status hierarchy. Third, every group has a religion, which is typically used to justify the status hierarchy and the existing moral and legal systems. This suggests that there are three overriding themes in individual lives: (a) efforts to get along with other people (because we live with them); (b) efforts to attain some power, status, and control of resources (more is always better); and (c) efforts to make some sense out of our lives (by interpreting them in terms of a quasi-philosophical system).

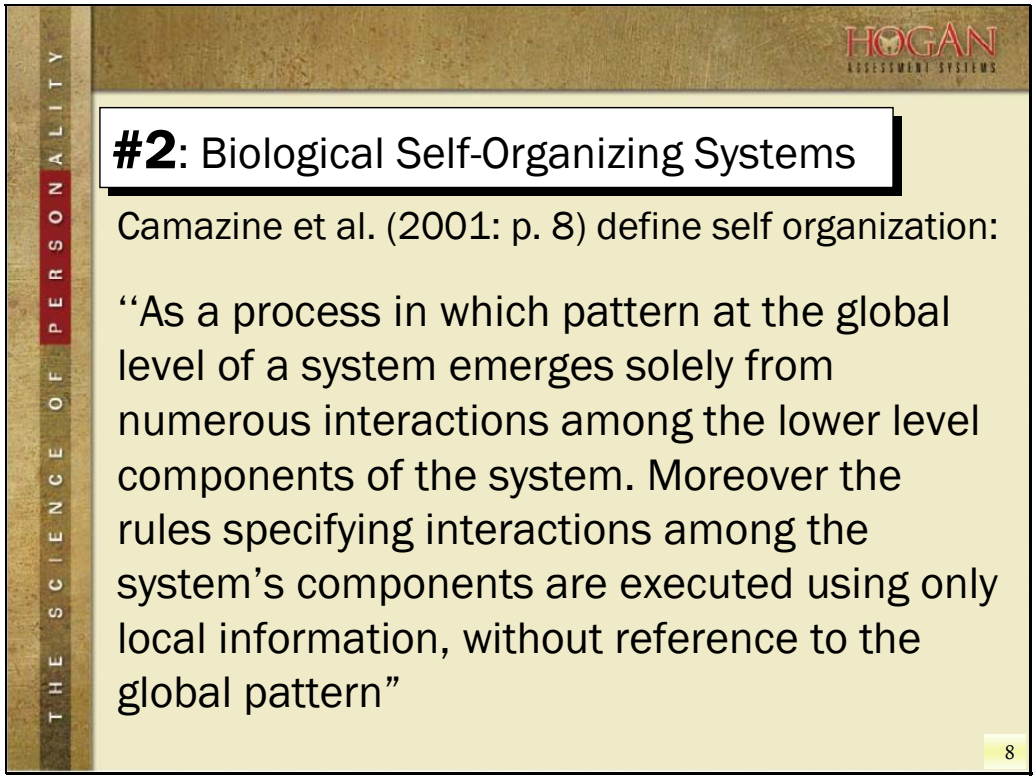
Personality psychology is also about individual differences. People differ from one another in many, many ways. These three generalizations—that people want acceptance, status, and meaning—suggest what the most important domains of individual differences might be. The first domain will concern individual differences in the desire for, and the ability to obtain, social acceptance and support. The second will concern individual differences in the desire for, and the ability to obtain, status, power, and the control of resources. The third will concern individual differences in the desire for meaning and purpose in life. I have suggested three important generalizations about human nature. I have suggested three important vectors of individual differences. Together, these point to a measurement agenda for personality psychology. In addition, please note that leaders are people who excel in their ability to gain acceptance and support, power and status, and to make meaning.

Finally, it is important to note that there is not one definition of personality, there are two. There is personality from the view of the actor, and personality from the view of the observer. Personality from the view of the actor—your view of you—is identity. Personality from the view of the observer—our view of you—is reputation. Identity and reputation are different, although somewhat related, concepts, and they have different implications for assessment. Self-reports—statements about who you think you are—are almost useless as data sources in and of themselves. As Freud might say, the you that you know is hardly worth knowing— because you made it all up. In any case, the study of identity is not very advanced, and has yielded few reliable generalizations.

On the other hand, reputation has several attractive features as a data source. First, we have a well-developed vocabulary for talking about reputations, and that is the vocabulary of trait words. Trait words are what we use to describe other people, and our descriptions of others are, in fact,

their reputations. Second, we have a well-developed taxonomy of trait words, and it is the so-called Five-factor Model. Recent research indicates that there are more than five dimensions of normal personality, but the point is that we have a well-defined structure in terms of which trait terms can be organized. Third, we can use trait words reliably to characterize others' reputations. And finally, reputations are immensely useful for predictive purposes. The best predictor of future behavior is past behavior, reputations are a summary of past behavior, therefore reputations are the best information we have about future behavior.

The measurement model that we use involves taking the statements that people make about themselves and then determining, in an empirical way, the links between their self-descriptions and their reputations. Take, for example, such statements as, "I can get this country moving again" or "I can take this company to the next level." Our research indicates that peers and subordinates describe people who say this as egotistical narcissists. The trick, then, is to take what people say about themselves and translate it into what other people say about them. Conversely, it is always risky to take what people say about themselves at face value; the data are quite clear that people are poor judges of how they are seen by others (cf. Harris & Schaubroeck, 1988; Mabe &West, 1982). When a manager tells me, "My staff respects me," I always say, "Can we take a vote"?



Well, this is where the rubber hits the road. Two quite independent substantive papers and a book have been published recently, whose summary of the human cognitive system seems to explain why we are now going nowhere fast in I/O psychology and individual differences in general.

The book is by Camazine, Deneubourg, Franks, Theraulaz, & Bonabeau (2001) Self-organization in biological systems. Now, we are not going to get bogged down in the wondrous new world of complexity theory (!)– it is sufficient for our purposes here to offer a simple textbook definition, a slightly longer explanatory piece, then quote from two recent papers whose authors explain most eloquently the problem for psychology with the 2nd Big Idea.

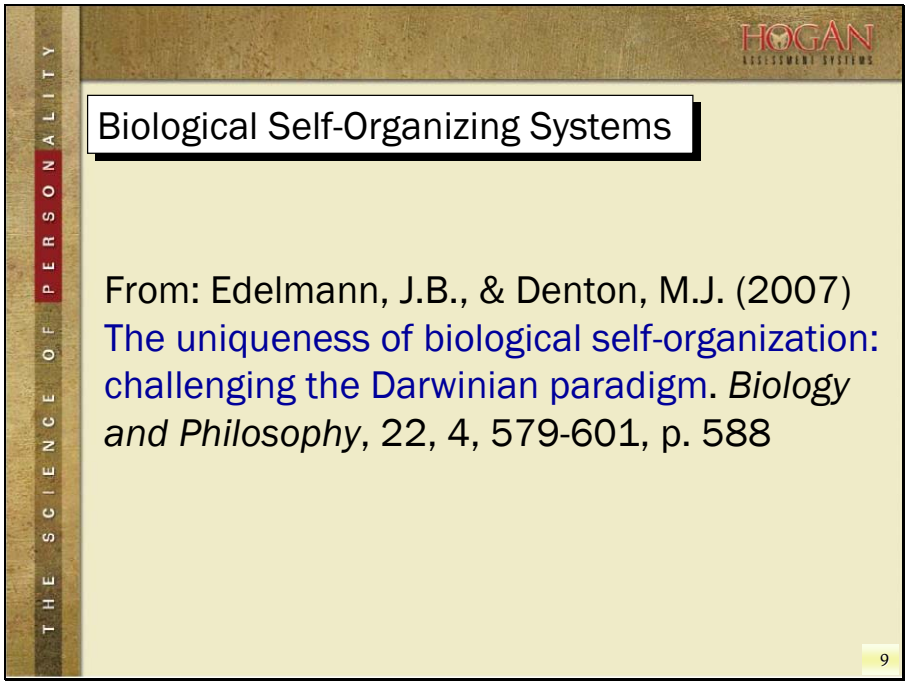
From Coveney and Highfield (1995) .. .p. 425 ..

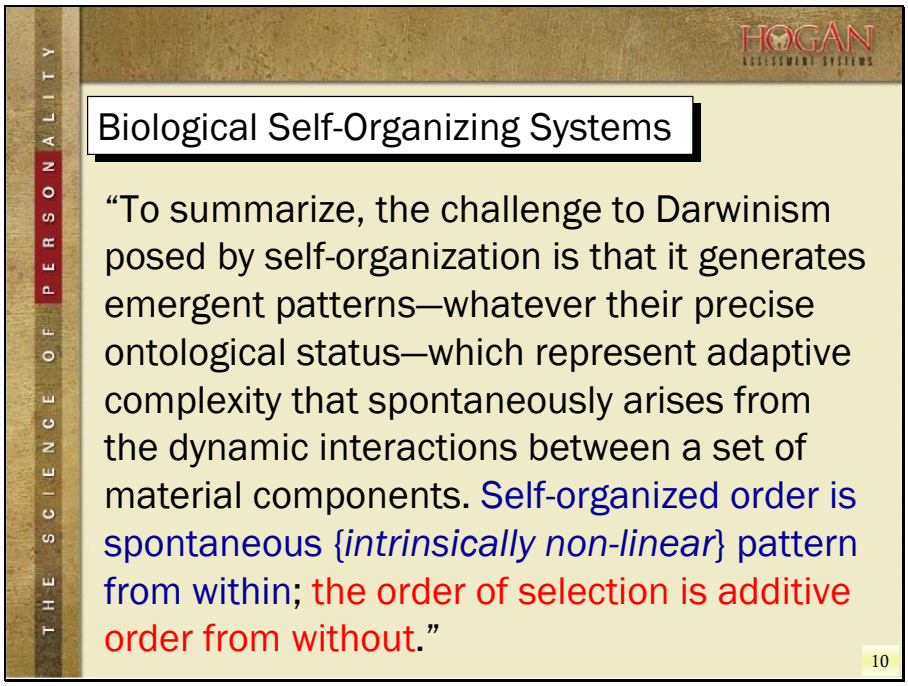
> *Complexity is the study of the behavior of macroscopic collections of simple units (e.g. atoms, molecules, bits, neurons) that are endowed with the potential to evolve in time.*

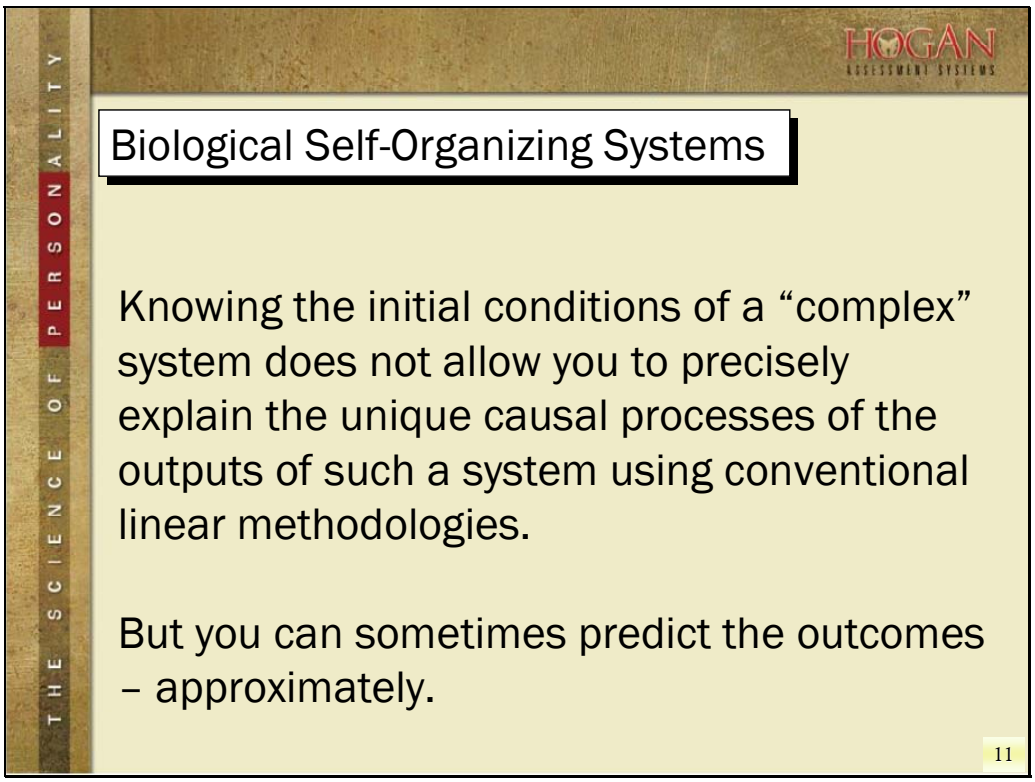
From Svyantek & Brown (2000), p. 69 …

> *Nonlinear systems theory proposes that the behavior of systems in the world is affected by many variables that interact strongly with each other, whereas more traditional systems theory assumes that a few variables, interacting weakly with each other, determine the behavior of systems (Liebovitch, 1998). The nonlinear view of systems used to describe* ***complex*** *systems' behavior originated in the physical sciences of physics and chemistry and is commonly known as chaos theory or* ***complexity theory****.*
>
> *Researchers use nonlinear methods to describe the behavior of complex systems (Gallagher & Appenzeller, 1999), which are those systems whose behaviors cannot be explained by breaking down the system into its component parts. Explaining the behavior of a complex system requires understanding (a) the variables determining the system's behavior, (b) the patterns of*

interconnections among these variables, and (c) the fact that these patterns, and the strengths associated with each interconnection, may vary depending on the time scale relevant for the behavior being studied. …







From: Edelmann, J.B., & Denton, M.J. (2007) The uniqueness of biological self-organization: challenging the Darwinian paradigm. *Biology and Philosophy*, 22, 4, 579-601.

I'm quoting these authors here to show that complexity theory has even challenged Darwinian Evolution as the sole mechanism for evolutionary adaptation and change. This is similar to the challenge facing behavior geneticists with the new knowledge of the existence and functions of microRNA (Clarke and Sanseau. 2007; The Economist, 2007). The pace of change in sciences related to psychology is frankly astounding.

pp 585-586.

> I*n all self-organizing systems the final emergent complex global form C (often referred to in the terminology of dynamic systems as the attractor of the system), arises from the spontaneous interactions of its components, a, b, c, and d (Camazine et al. 2001: 32). In the case of self-organization the form and properties of C are not built up bit-by-bit gradually by the simple linear addition of successive incremental steps—i.e., of a + b + c + d—but arise out of complex non-linear interactions (Strogatz 1994: 9) involving various types of self-reference and feedback (Gleick 1998: 167–68; Gribben 2005: 79). Being genuinely emergent, C is more than the linear sum of its parts. Rather, we may more properly represent C as a*b*c*d. The process is more akin to a sudden phase transition (Kauffman 2000: 35; Florey 1969; Haken 1977) rather than any sort of incremental stepwise process. The unexpected emergence of self-organized complexity has been described as ''magic'' by Kauffman (2000: 35, see also Lewin 1992: ch. 2) and something of its ''magical'' quality is illustrated by the complex emergent patterns—sometimes fantastically intricate— generated by cellular automata programmed to follow remarkably simple rules (Levy 1993; Wolfram 2002). Cumulative selection (Darwin 1872; Dawkins 1986; Dennett 1995) is therefore a completely different and far more prosaic mechanism—as far from something like a phase transition as can be imagined. The realm that selection rules is a predictable realm characterized by intermediates, continuity and gradual linear additive functional transitions. On the other hand, the realm of self-organized complexity is*

an unpredictable realm of sudden spontaneous emergent complexity that is generated by non-linear interactions via something like a phase transition. This is a realm where saltation, emergence, spontaneous sudden change and bifurcations rule; a realm in which the concepts of intermediacy, gradualism and continuity, so central to the Darwinism, no longer apply. In Misteli's words (2001: 183):
''Typically self-organized systems undergo transitions which occur suddenly and typically without intermediates. Self-organized systems are frequently in a state of criticality; that is, a point at which systems properties can undergo sudden change. Transitions between distinct assembly patterns are not gradual but sharp.''

*You can assemble a LEGO toy bit-by-bit through cumulative selection, but you cannot gradually cross bit-by-bit the ''phase transition'' that separates the pre-organized constituents of a self-organized system from their post self-organized newly emergent form. The ''good design'' (Dennett 1995: 144) generated by selection is ''hard earned'' over time; self-organization is unearned ''for free'' (Kauffman 1993). Generating adaptive complexity by cumulative selection is climbing ''mount improbable'' gradually in hard earned steps over perhaps thousands of generations (Dawkins 1996). Generating adaptive complexity by self-organization is climbing mount improbable in one saltational step! Cumulative selection may gradually put together or fine tune the physical properties of each of the individual constituents a, b, c, and d of a self-organizing ensemble, but not the collective property of the ensemble (a*b*c*d) to self-organize into the emergent higher order structure C. This collective property depends on the temporal and spatial co-presence of all the constituents of the ensemble in a particular physiological environment. For this reason classic cumulative selection can never be the determining cause of the emergent complexity (or indeed of any emergent properties) generated by self organization in any complex system.*

Ok – bear with us for one more quote – this one entirely focused on behavior, as it is probably the final essential article in understanding the real implications of this complex systems theory.

Richters, J.E. (1997) … The Abstract to the paper …

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*

The information we need is associated with those two concepts – equifinal and multifinal.

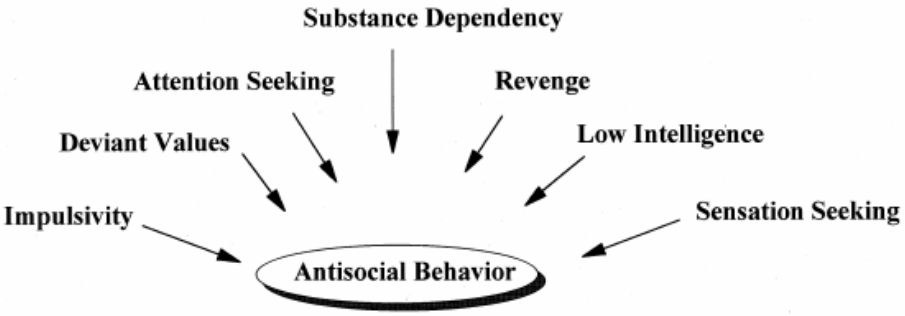


Figure 1. Equifinality in human functioning as an illustration of different personality structures and processes that might give rise equifinally to persistent antisocial behavior in different individuals, depending on a matrix of other internal and external factors.

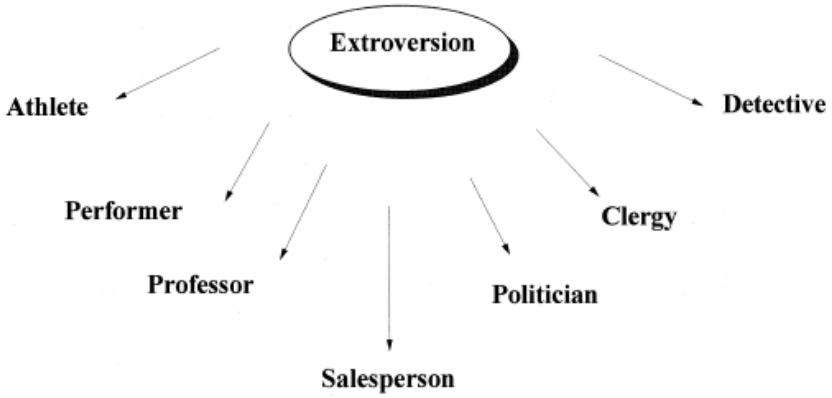


Figure 2. Multifinality in human functioning as an illustration of different overt behavior patterns that might arise multifinally from a common characteristic such as extroversion, depending on a matrix of other factors internal and external to an individual.

Equifinality in open systems.

All living organisms fall into the category of open systems, which exchange energy and matter with the environment and are not therefore subject to entropy. Beyond thermodynamic openness, living organisms are distinguished from closed systems by unique functioning characteristics such as growth, adaptiveness, regeneration, emergence, differentiation, dynamic self-stabilization, and self-reorganization. Organisms are also distinguished from closed systems by the functioning capacities of equifinality and multifinality.

Equifinality *is the ability to reach similar outcomes or end states from different starting points and through different processes (Figure 1). Conversely,* ***multifinality*** *(Figure 2) is the ability to reach different outcomes from similar start points and/or through similar processes (von Bertalanffy, 1968). Thus, in contrast to closed systems, the initial conditions of open equifinal systems imply nothing about their end states and their end states imply nothing about their initial conditions.*

Now, let's sit back for a moment and digest what we've been saying:

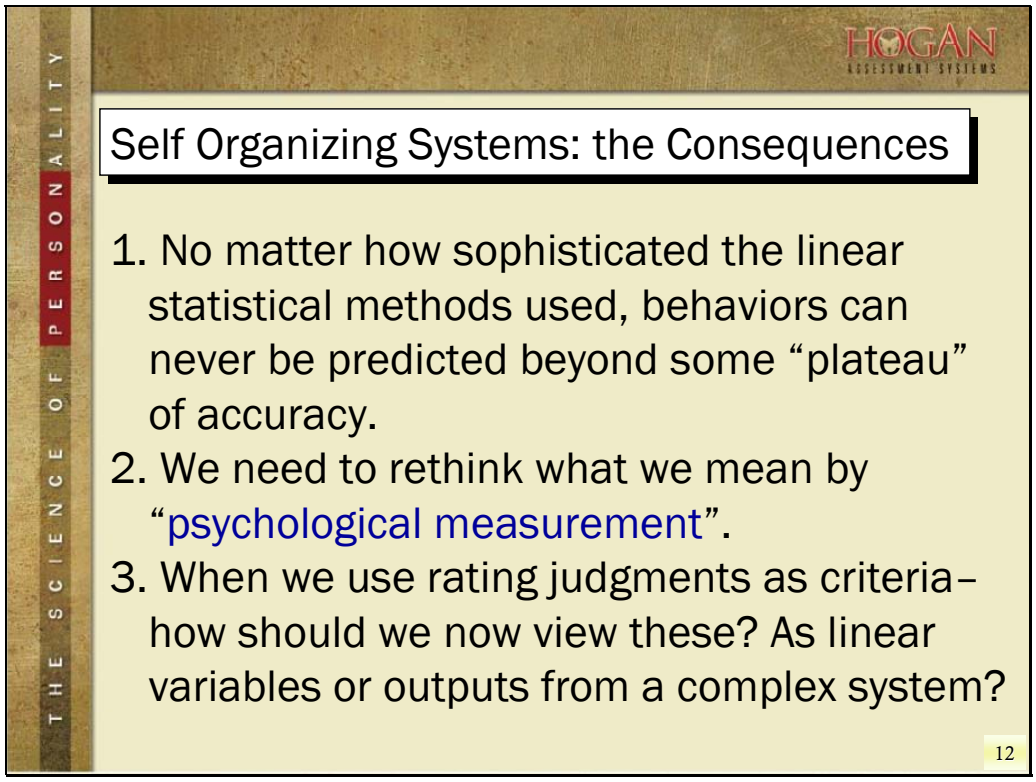
1. The brain is a complex, open system.
2. We now know such systems cannot be described by conventional closed-form mathematics.
3. Complex systems produce "emergent" properties.
4. These emergent properties are the result of nonlinear interactions between many "units" and even sub-systems.
5. Untangling cause within systems exhibiting equifinal and multifinal properties is beyond current conventional analysis methodologies and may never be possible.
6. Assuming behavior and cognitions can be "measured" as one might measure length looks ridiculous.

So why do psychologists still insist on working with simple statistical models such as linear regressions, Structural Equation Models, IRT theory, and psychometric models when the variables we are dealing with look intrinsically nonlinear, indeed "complex" in terms of all that means..

The last 80 years have been a valiant effort to model human attributes and behaviors as though they were "closed" systems. It worked, up to a point. That point was probably reached some 30 years ago.

This is not a call to cease using quantitative statistical methods, but to note that heuristic and dynamical systems algorithms are probably going to work much better for us.

And in this new world, it is theory-driven and robust predictive accuracy which will rule.



I think we still have two options:

A) hope that increasingly complex measurement models, mathematical and statistical methodologies will take us to the next level of predictive accuracy and understanding of psychological phenomena. That is where the next generation of quantitative psychologists will attempt to take us.

Or

B) Think again, and figure out how to work with psychological variables and behaviors which represent the output from complex systems. i.e. rethink, retool, and build upon the foundation of the earlier work. That is where some of us are headed.

Quite frankly, Joel Michell (1997) gave us the final wake-up call that something was dreadfully wrong with psychometrics and psychology's view of what constitutes quantitative measurement.

In 2007, where are we? Still plodding away with the same kinds of highly constrained data models for the closed, mechanistic psychological constructs we've always worked with – still observing validity coefficients of 0.2/0.3 (if lucky), with our innovation largely confined to new kinds of statistical data models and technology for delivering questionnaires.

It is no wonder statisticians like David Freedman (2005, p. 200) note:

> *In the social and behavioral sciences, far-reaching claims are often made for the superiority of advanced quantitative methods – by those who manage to ignore the far-reaching assumptions behind the models*.

BUT

There are no snappy answers here. As Richters (1997) notes, a paradigm change in thinking is required. This will not happen until results from those, like ourselves, show that greater predictive accuracy and financial benefits can be gained from a more imaginative and innovative approach to theory construction and data analysis.

The Hogans have already begun this process with socioanalytic theory and their stance on traits. Likewise the development of the model of "Good Judgment" and the Hogan Business Reasoning Inventory

The second step is to reconsider what we call "error", reliability, internal consistency, and validity, in a world where measurement is no longer strictly quantitative, and data distributions not necessarily normal. Engineering has robust definitions of reliability that do not rely upon hypothetical true scores or "item universes". Measurement equivalence also takes on an interesting perspective – in a world where the variable being assessed is no longer the equivalent to a measure of length or weight.

But then as some commercial companies are now coming to realize, you don't need to use questionnaires anymore to acquire reliable self-report rating data (www.monster.com, www.iyomu.com). This move toward single–item assessment is growing, especially in epidemiological, quality of life, and marketing (Bergkvist, & Rossiter, J.R. (2007)).

I think what disturbs many of us is the realization that quantitative measurement of all psychological attributes, as has been achieved in physics can in principle never be attained. Peter Schonemann said as much in 1994, and explained why. This should be a source of challenge and opportunity, not depression.

I hope that some reading this will grasp the fact that the system under examination (our brain) is capable of generating broad regularities in behaviors, but is by definition, complex, adaptive, self-organizing, dynamical, and non-linear in function. We can approximate its outputs to some degree, and our contention is that if you drop the insistence of assuming every set of observations conforms to a particular kind of data model, and work with the reality of variables which are virtually by necessity

"fuzzy", then we can indeed expect to improve our predictive accuracies, because we stop trying to predict exact numerical values and look instead to working with classes and orders; the way we work with most judgments and ratings in the real-world. There is nothing to stop us trying to predict exact numerical values "as usual" – just don't stop there if it doesn't work out, as there are good reasons not to expect this to work.

The Bottom Line

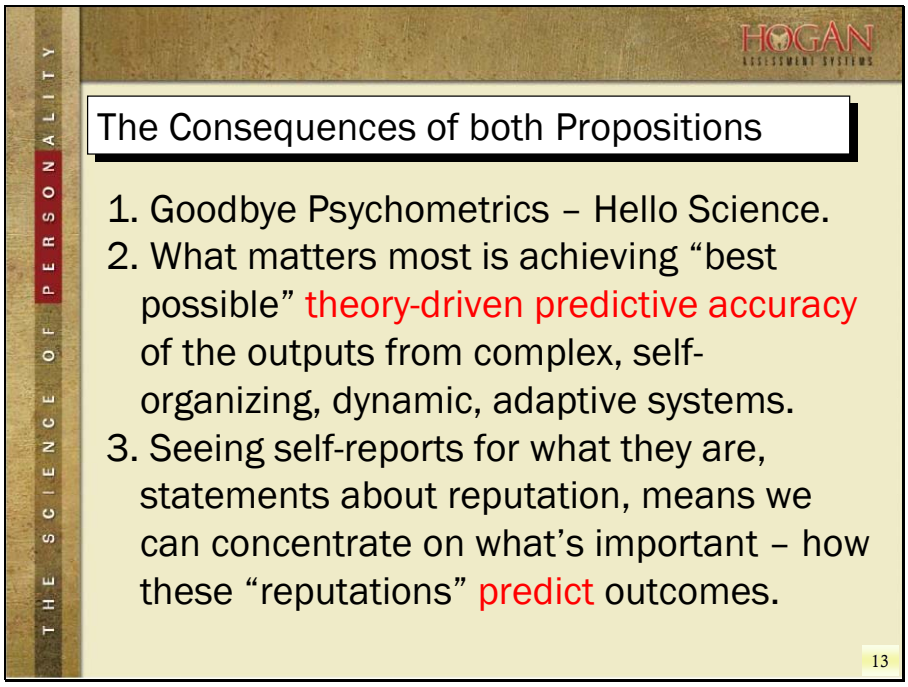
In the end, for clients and users of psychological knowledge, what is at stake is a magnitude shift upwards in predictive accuracy, simply because we have a more realistic view of the constructs we are attempting to assess and predict.

Clearly, this improvement is not coming anytime soon with IRT, SEM, HLM, MCMC, or all the other "alphabet-soup" methods that now proliferate our journals. Instead, we look at the data itself rather than models of it, and likewise with the properties of criteria to be predicted. Predictive accuracy of meaningful criteria is the goal, not correlations, not t-tests, and not arcane structural covariance models.

The statistics or summary data parameters we might use are designed to find exactly that which the client desires, if it exists, in the data. But, we also now have available the freedom to innovate theory-driven algorithms to pick their way through datasets to find us the best solutions. We know we're not dealing with linear generating systems– so we stop pretending that true scores exist, or that judgments/ratings must conform to equal-interval or even remotely linear functions.

And don't think that this presentation is anything but the brow-beating of two rather obvious (in hindsight) propositions. The Big Ideas are simple to understand – once you stop to think about matters.

But, the consequences are substantive. We are exploring our own way through these – as are other scientists who are beginning to work at the same paradigm threshold.



THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

A closing statement

From David Freedman (2003) Statistical Models: Theory and Practice. Cambridge University Press, p. 200 …

"The goal of empirical research is – or should be – to increase our understanding of the phenomena, rather than displaying our mastery of technique".

14

Finally, I've included a small example of how assumptions about data color the way we might approach their analysis.

One way many consultants and psychologists approach isolating variables which might be used for selection purposes is to compute the correlations between the variables, looking for those correlations which exceed some threshold (usually 0.2 or so) – and then proceed to recommend these (or do further analysis with them) for selection purposes. They might also use multiple linear regression to discover those variables which are the best predictors.

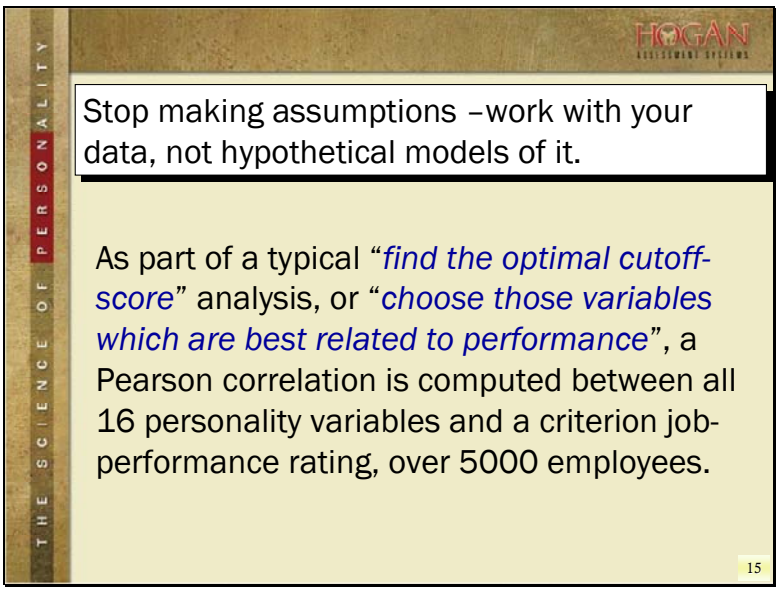
The problem with this is that the implicit data model under which they are operating is that the predictor variables and criterion are continuous-valued and normally-distributed. **What if they are not**, and what if such non-normality is not serious (skewness and kurtosis < 0.3).

The admittedly "manufactured" dataset I used, of 5000 cases, shows how badly one can get it wrong when failing to use methods which are sensitive to the actual data collected, rather than some idealized data-model view of it. The correlation between two variables, Adjustment and Job Performance rating, was just 0.16. So it would be rejected as "not worth bothering with". Yet, by using a simple iterative cutoff-score algorithm, it was shown that overall predictive accuracy was at 90%, given a dichotomized criterion rather than the usual "equal-interval-scaled" variable commonly used in these procedures.

All the algorithm did was compute the numbers of cases predicted correctly and incorrectly when each score value of adjustment was used as a cut-off score, and display the prediction accuracies graphically as an optimization graph. We could have overlaid financial costs and benefits as a secondary graph, or even placed constraints on the optimization to yield the optimal financially rewarding solution, taking into account selection ratios. All of this is a long way from a simple correlation coefficient.

Essentially what we did was to treat the Adjustment score as ordered classes, with an outcome reduced to two classes – accept-reject.

This example is not a blue-print for any further analysis – it is here to show how working directly with data and constructs reveals outcomes which would never even be seen by those working with "text-book" data-model statistics. And yes, these algorithmic methods owe more to statistical data mining and machine learning than "psychological statistics".



THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

- Those variables correlating above 0.25 are considered potential variables for further cutoff-score optimization/regression analysis
- A correlation between Adjustment and Job Performance is just 0.16. There is no range restriction on either variable.
- The variable is discarded for cutoff score analysis as the correlation is considered too low for the variable to be of much predictive value.

16

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

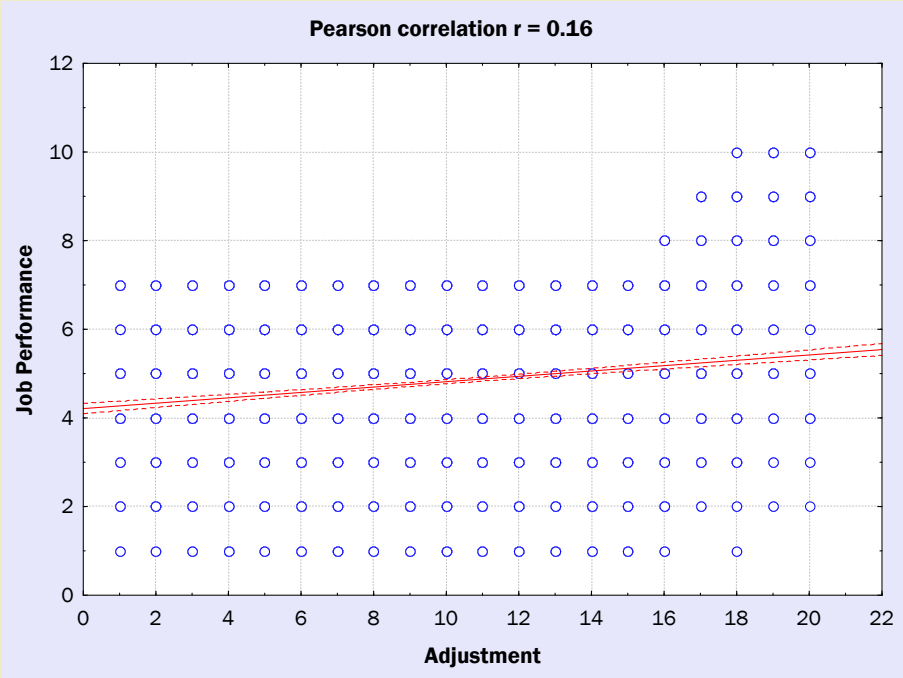
- But, let's get clever. We ask the employer to put the "flesh on the bones" of that performance rating.
- What rating value (1-10) do the supervisors feel is indicative of really "stand-out", "if only we could get more of these" employees?
- They say 8 and above.
- So, we we dichotomize our performance scores; 0=reject, 1=select *(performance score 8 and above)*

17

THE SCIENCE OF PERSONALITY

HOGAN ASSESSMENT SYSTEMS

- who said I/O data must always be bivariate normal?
- Why should performance ratings/judgments be linear?



18

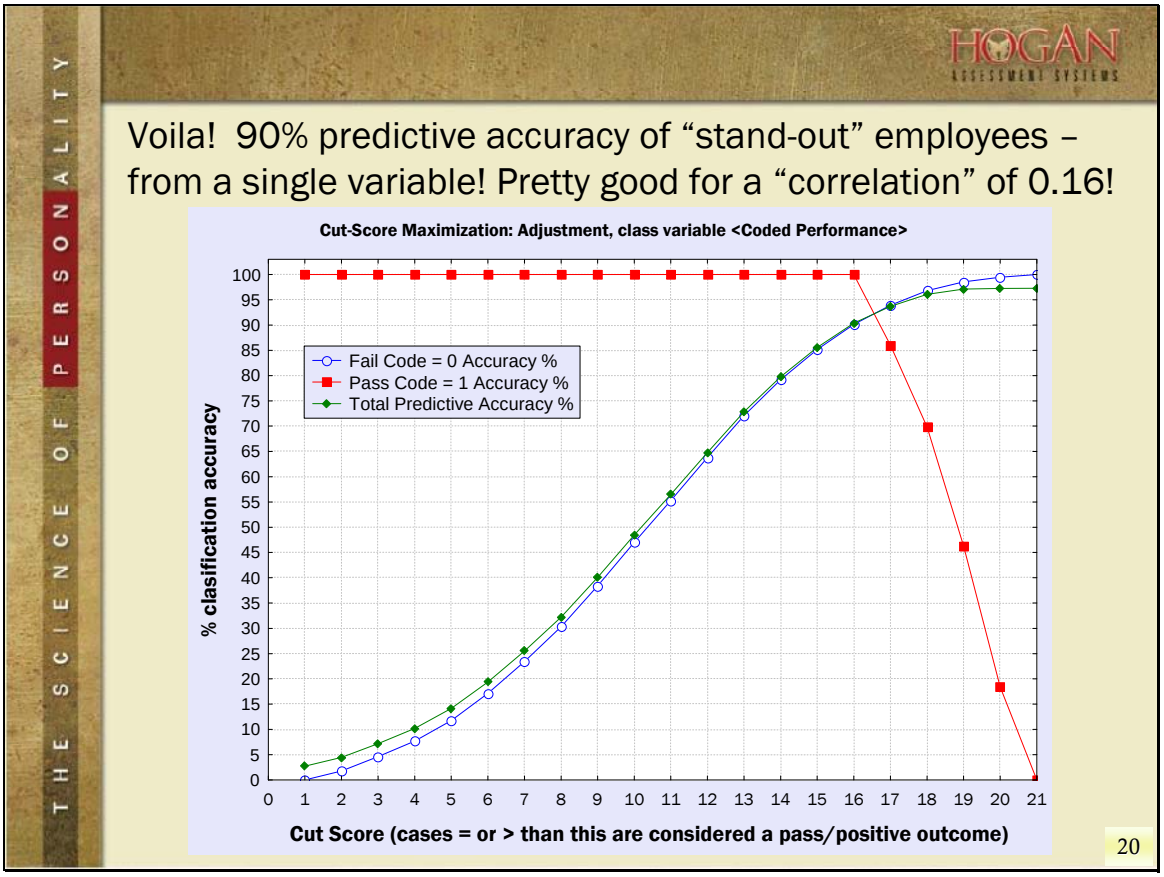
THE SCIENCE OF PERSONALITY

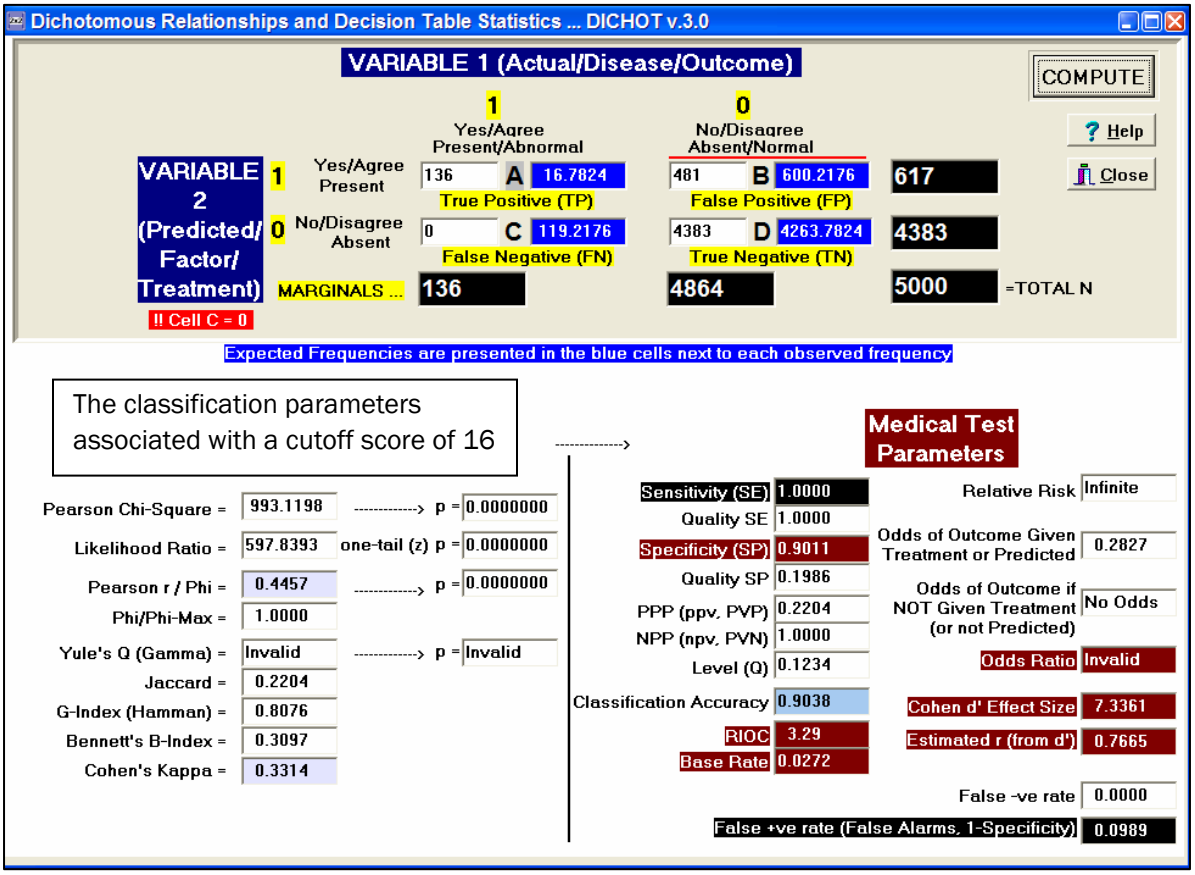
HOGAN ASSESSMENT SYSTEMS

- We forego statistical data models (i.e. data are sampled from a theoretical bivariate normal distribution) – and just "go figure" exactly what the employer wants us to do – maximize the predictive accuracy for an employer-specified criterion.
- We attempt to predict the kind of outcome the employer wants – **using any means possible**.
- We let the data speak – not the statistical model.
- Because we are looking at developing a cutoff-score - we might look at each predictor variable in turn to see how a cutoff-score placed on it would predict the dichotomized performance variable…

19

And the final optimization result … suggesting a cutoff-score of 16 and above …





References

Bergkvist, L., & Rossiter, J.R. (2007) The predictive validity of multiple-item vs single-item measures of the same constructs. *Journal of Marketing Research*, 44, 2, 175-184.

Borsboom, D. (2006) The attack of the psychometricians. *Psychometrika*, 71, 3, 425-440.

Camazine S, Deneubourg J-L, Franks NR, Theraulaz G, & Bonabeau E (2001) Self-organization in biological systems. Princeton University Press, Princeton

Cherneyshenko, O.S., Stark, S., Drasgow, F., & Roberts, B.W. (2007) Constructing personality scales under the assumption of an Ideal Point response process: toward increasing the flexibility of personality measures. *Psychological Assessment*, 19, 1, 88-106.

Clarke, N.J. and Sanseau, P. (2007) *MicroRNAs: Biology, Function, and Expression*. Eagleville, PA: DNA Press.

Coveney, P. & Highfield, R. (1995) *Frontiers of Complexity: The Search for Order in a Chaotic World*. New York: Faber and Faber.

Edelmann, J.B., & Denton, M.J. (2007) The uniqueness of biological self-organization: challenging the Darwinian paradigm. *Biology and Philosophy*, 22, 4, 579-601.

Freedman, D.A. (2005) *Statistical Models: Theory and Practice*. Cambridge, UK: Cambridge University Press.

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

Garlick, D. (2002) Understanding the Nature of the General Factor of Intelligence: The Role of Individual Differences in Neural Plasticity as an Explanatory Mechanism. *Psychological Review,* 109, 1, 116-136.

Jackson, J.S.H., & Maraun, M. (1996) The conceptual validity of empirical scale construction: the case of the Sensation Seeking Scale. *Personality and Individual Differences*, 21, 1, 103-110.

Jackson, J.S.H., Maraun, M. (1996) Whereof one cannot speak, thereof one must remain silent. *Personality and Individual Differences*, 21, 1, 115-118.

Maraun, M.D. (1997) Appearance and Reality: Is the Big Five the Structure of Trait Descriptors? *Personality and Individual Differences*, 22, 5, 629-647.

Maraun, M.D. (1998) Measurement as a Normative Practice: Implications of Wittgenstein's Philosophy for Measurement in Psychology. Theory & Psychology, 8(4), , 435-461.

Michell, J. (1997) Quantitative science and the definition of measurement in Psychology. *British Journal of Psychology*, 88, 3, 355-383.

Michell, J. (2004) Item Response Models, pathological science, and the shape of error. *Theory and Psychology*, 14, 1, 121-129.

Revelle, W. (In press) *The contribution of Reinforcement Sensitivity Theory to Personality Theory*. In Phillip Corr (Eds.). Reinforcement Sensitivity Theory of Personality. Oxford University Press

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

Schonemann, P. (1994) *Measurement: the Reasonable Ineffectiveness of Mathematics in the Social Sciences*. In I. Borg and P.Mohler (Eds.). Trends and Perspectives in Empirical Social Research. Walter de Gruyter.

Smillie, L.D., Pickering, A.D., & Jackson, C.J. (2006) The new Reinforcement Sensitivity Theory: implications for personality measurement. *Personality and Social Psychology Review*, 10, 4, 320-335.

Svyantek, D., & Brown, L. (2000) A Complex - Systems Approach to Organisations. *Current Directions in Psychological Science*, 9, 2, 69-74.

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

Zuckerman, M. (1996) "Conceptual Clarification" or confusion in "The study of Sensation Seeking" by Jackson and Maraun. *Personality and Individual Differences*, 21, 1, 111-114.