World Futures, 64: 383–398, 2008

ISSN 0260-4027 print / 1556-1844 online
DOI: 10.1080/02604020802301212

Routledge
Taylor & Francis Group

TOWARD A CROSS-SPECIES MEASURE OF GENERAL INTELLIGENCE

MICHAEL LAMPORT COMMONS
Department of Psychiatry, Harvard Medical School, Cambridge, Massachusetts, USA

SARA NORA ROSS
Dare Institute and ARINA, Inc., Bethel, Ohio, USA

Science requires postformal capabilities to compare competing explanations and conceptualize how to coordinate or integrate them. With conflicts thus reconciled, science advances. The Model of Hierarchical Complexity facilitates the coordination of current arguments about intelligence. A cross-species measurement theory of comparative cognition is proposed. It has potential to overcome the lack of a general measurement theory for the science of comparative cognition, and the lack of domain-general mechanisms for evolutionary psychologists. The hierarchical complexity of concepts and debates as well as the new theory are scored, and demonstrate the postformal hierarchical complexity of the proposed theory.

KEYWORDS: Comparative cognition, domains, evolutionary psychology, hierarchical complexity, g, intelligence, IQ, measurement theory.

The primary purpose of this article is to use the universality of the Model of Hierarchical Complexity to respond to an identified scientific need. The response is the proposal of a cross-species, cross-domain measurement theory that reconciles long-standing debates and bridges conceptual divides in studying human and non-human animals. In the context of this *World Futures* issue, another purpose is to demonstrate the kind of contribution possible with Metasystematic stage 12 approaches to address conceptual and scientific challenges that have proved difficult otherwise.

Introductory comments about studies of comparative cognition and evolutionary psychology and the objectives of the article begin the discussion. Next, the concepts, background, and key debates in human and animal cognition are introduced, along with several applied examples of hierarchical complexity scoring of animal tasks. Then, the proposed measurement theory, its multiple indices, and how to calculate them are described. Before concluding, an analytical section offers a hierarchical complexity-based interpretation of several concepts in use, and

Address correspondence to Michael Lamport Commons, Ph.D., Dare Institute, 234 Huron Ave., Cambridge, MA 02138-1328, USA. E-mail: commons@tiac.net

scores both of the intelligence debates discussed—modularity and domain-general mechanisms—as well as the measurement theory proposed here to reconcile them.

COMPARATIVE COGNITION AND EVOLUTIONARY PSYCHOLOGY

The branch of science called comparative cognition describes itself as the study of animal intelligence,[1] that is, animals' mechanisms and evolution of cognition.[2] "Animals" in this colloquial usage exclude human animals. This indicates an artificial divide of the scientific Kingdom *Animalia,* which by definition includes humans. Comparative cognition is currently comparing non-human animals and excluding human animals. Thus a conceptual issue stands at least partly in the way of a science of genuinely comparative cognition across animal species.[3] This may rest upon assumptions that "animal" and human cognition are incomparable.

In a similar vein, the founders of evolutionary psychology—which they call a way of thinking about psychology—describe its goal as research into the human mind's design (Cosmides and Tooby, 2007). Their work and that of others (e.g., Thagard, 2006) indicates the lack of a dividing line between emotion and cognition. Human emotion is not a factor that may separate humans from other members of *Animalia.* What basis exists for comparative cognition to not apply to all members of the animal kingdom?

Even though current theories have generally maintained this human–other animal divide in the study of cognition, some early studies of intelligence considered whether or not animals could think (Romanes, 1888). Darwin's (1871) conclusion in the fourth chapter of *Descent of Man* noted that the difference in human and animal mind was "certainly" one of degree, not one of kind. Thorndike's work supported that notion. Doing work with both humans and other animals, he developed basic laws of learning in animal behavior, the Law of Effect (Thorndike, 1898, 1911). It states that responses to a situation that are followed by satisfying events are strengthened and responses that are followed by unpleasant ones are weakened.

This article supports dissolving unnecessary conceptual and methodological divides. The current proposal can serve to bridge two divides that have prevailed as well as address another void. We propose that a science of comparative cognition needs such a theory to enable comparisons of the performances of different species of animals and different groups of people. One situation to address is that current theories based on humans may not readily apply to other species. Another is the controversy over whether cross-cultural comparisons of human groups can be made, and if so, whether they can use current theories and tests developed within one culture, for instance. Finally, the need of evolutionary psychologists for mechanisms that apply across domains, that is, domain-general, has been recognized (Geary, 2004). To respond to these needs, the proposal outlined here is for a cross-species, cross-domain measurement theory of cognitive performance.

CONCEPTS AND DEBATES

The Concept of g, General Intelligence Factor in Humans

The concept of a general intelligence factor, g, has surfaced across disparate researchers' work for over a century. A prevalent although not universally accepted concept, the fact of its utility in accounting for test results is significant, and is the point emphasized here. While referred to as a general factor, it is also significant to note that researchers have recognized additional factors operating in tandem with it.

Sir Francis Galton (1869, 1892, 1962) originated the concept of intelligence and studied its heritability. Many studies have shown that g is at least 50 percent heritable and thus, can be passed down from generation to generation (DiLalla, 2000). Intelligence testing, although empirically driven, has been based on a vague understanding of the nature of intelligence and of domains of human endeavor it shows up in. Spearman (1904) explained positive correlations in studying school children's performance in terms of a dominant factor of g for "general" intelligence. His model relied on two factors: g, governing performance on all cognitive tasks, and a factor specific to an individual mental task, where individual abilities would make one person more skilled at one cognitive task than another. Later factor analyses would indicate more factors were involved.

Thorndike and colleagues (1904) distinguished three broad categories of intellectual functioning: abstract, mechanical and spatial, and social. Later, they developed an intelligence test to measure intellect on an absolute scale (Thorndike, 1920, 1927) using a test design logic that modern intelligence tests would later use. Also recognizing the disparate classes of activity in human intellect, Thurstone (1931) introduced factor analysis to measure relationships among many variables. It allows numerous intercorrelated variables to be reduced to fewer dimensions, called factors. This enables detection of structure in the relationships between variables so they can be classified. The factor that persists today is g, the first factor of seven named by Thurstone. The accumulation of "cognitive" testing data and improvements in analytical techniques have preserved g's central role and led to the modern conception of g (Carroll, 1993). Skottke (March, 2006) argues that general intelligence, g, can be described as the ability of an individual to acquire and apply knowledge. A hierarchy of factors, with g at its apex and group factors at successively lower levels, is a widely, albeit not fully, accepted model of cognitive ability. Extensive critiques on conceptual and methodological grounds exist, but further discussion is not included here because it is beyond the scope set for this article. For an introductory summary, however, see Shalizi (2007).

A Standardized IQ Test and Hierarchical Complexity

Despite the conceptual, methodological, and even ideological arguments attending g, we do not dismiss it out of hand. The premise here is that once the Model of Hierarchical Complexity and domains are coordinated with the notion of differences in human intelligence, that a sufficiently complex understanding of a dimension akin to g is not only possible, but approximately measurable. A current project that we are working on represents such coordination. Test items and their correct responses in several standardized IQ tests are being scored for the stage of hierarchical complexity they represent. All items are able to be scored using the Hierarchical Complexity Scoring System (HCSS) (Commons et al., 2007) because they represent tasks. An empirical example of the relationship between a standardized IQ test and orders of hierarchical complexity is provided for the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Table 1). This test is designed for U.S. children from ages of about 4 to 6½ years. Professional ethics require that the content of its items and correct responses to the items are not shown; only a small selection of items is given for further masking of the proprietary information. Note that in comparison to the age range of children to whom the test is administered, performances at Primary stage 7 would be common in children from 6–8 years of age, and Concrete stage 8 would be common at 8–10 years of age.

When the content of the specific tasks is eliminated, as already mentioned, and only the general nature of the task remains as the descriptor, one might consider if

Table 1
Scoring of Selected (Masked) Tasks in the WPPSI

Tasks Represented by Selected WPPSE Test Questions	Stage of Performance	Score Explanation
Say the name of an indicated object.	4 Nominal	Relate concepts to others through a word; single words
Perform the instructed action.	5 Sentential	Correct use of pronouns, sequences actions and nominal concepts in subject, verb, and object
Indicate how many objects of a particular kind are on another object.	6 Pre-operational 5.8 Sentential transition	Count objects; combine numbers and simple propositions Smash transition step if verbal and physical answers do not match but one of them is correct
Say the kind of contents typically stored in a particular kind of container.	6 Pre-operational 5.8 Sentential transition	Make simple deductions Smash transition step if over generalization about container
Identify the last action to take before a particular task is accomplished.	6 Pre-operational	Organize sequential actions
Count how many specific parts make up the indicated object's value.	7 Primary	Simple arithmetic and concepts applied to concrete objects
Say the four things related to a particular thing everyone experiences.	7 Primary	Empirical rules involving time sequence
Say the name of the direction where a regular natural event happens.	8 Concrete	Relations among specified times and places

there are merits to aggregating such generic tasks and calculating a general level of intelligence. This example of the scorability of IQ-measurement approaches indicates not only the sorts of tasks included in one IQ test for children, but also the flexibility of the HCSS to score either content-laden material or content-free task descriptions. Its unidimensionality can thus bypass arguments referred to earlier about factor-based and other analyses, depending on applications.

Contrasting Arguments: Modularity and Domain-General Mechanisms

Evolutionary psychology has been concerned with the adaptiveness of the broad spectrum of human behaviors in the many domains in which humans function. This led understandably to a debate between two concepts: the modularity of intelligence (Cosmides and Tooby, 1994) versus general intelligence. The modularity concept in cognitive science conceives the mind as composed of independent, closed, domain-specific processing modules, that is, systems. The concept can be extended beyond the concept of "mind," however. Sperber (2002) asserts modularity's application to any biological mechanism. Modularity can be considered at five levels: morphological or architectural, developmental, neurological, genetic, and evolutionary (see Commons, 2006, for elaboration of these levels). In the modularity view, much modern human psychological activity is rooted in adaptations that occurred earlier in human evolution, when natural selection was forming the modern human species. Miller (2000) argues that one of evolutionary psychology's most distinctive ideas is the expectation that the control of human activity is massively modular. The brain evolved so that it is composed of hundreds of distinct psychological adaptations that evolved to solve distinct ancestral problems of survival and reproduction. Modularity is obvious for morphology: animals have distinct limbs, senses, and organs to do different things. Psychological and behavioral control modularity has been less obvious to psychologists, but evolutionary considerations of functional efficiency suggest the brain and behavioral control should be at least as modular as the body. The modularity view in studies of animal intelligence has generally prevailed because most animals show little sign or degree of general intelligence (Locurto, 2004). Critiques of certain strident dimensions of evolutionary psychology and contrasts with other social science approaches are summarized in Gintis (2007), for example, and not further discussed here.

The debate is whether g is domain-specific, as in modularity, or domain-general. Theories that posit domain-general processing assume mental activity that is distributed across the brain and cannot be separated into modular units either by conceptual abstractions or the physiological elements involved in any organism. Serious philosophical, theoretical, and methodological problems trying to localize cognitive processes in the brain have been identified (Uttal, 2003), one being that a successful taxonomy of mental processes has yet to be developed. Geary (2004) has argued that domain-general mechanisms are essential for evolutionary psychology because the human motivational, affective, behavioral, and cognitive systems have evolved to process social and ecological information (e.g., facial expressions) that covaried with survival or reproductive options during human evolution. Survival and reproduction systems involved gaining access to and controlling primary resources: the social (e.g., mates), biological (e.g., food), and physical (e.g., territory). Darwin's concept of natural selection as a "struggle for existence" applies to humans' struggle with other humans. This supports the argument for domain-general because such activities required integrating any modular brain and cognitive systems such as language with those brain and cognitive systems that support general intelligence. Geary makes a strong interdisciplinary argument for the domain-general mechanism view, and Chiappe and MacDonald (2005) argue the need for such a notion to understand human evolution.

Correspondence of Modularity with Domains

Despite the foregoing debates, the question of modularity or domain-general mechanisms of cognition is not one that has to remain in endless debate. Because values of g may have increased over evolutionary time as new organisms developed, there is a great need for defining g in a way that captures the issue of modularity versus generality in a systematic way. Modules are similar to the notion of domains. These modules are thought to be related to brain function that is specialized for tasks in a given domain.

Although there are currently no standard ways to define domains, one can rely on existing research to enumerate domains. It would seem that as the Model of Hierarchical Complexity is used to score more tasks, that in the process more domains may be discriminated. For example, many animals exhibit certain problem-solving behaviors in pursuit of food, and different behaviors in pursuit of mates and/or reproduction. Animals also have different kinds of interaction behaviors with others of their own species. Some animals pair bond, some live in social groups, and some are loners, coming together only to mate.

To develop a measure of g for animals will ultimately need taxonomies of discrete domains and the collection of tasks within those domains that can be performed by various animals. Such taxonomies have not yet been developed, nor are domains yet situated on any scale. At present, domains are referred to by names only, that is, they are nominal. From the animal literature, the domains for most animals are discrete. As a taxonomy of human domains is developed, we may note cross-domain tasks are inherent in many of them, for example, language and literacy. Major animal domains previously listed (Commons, 2006) are repeated later in the article. Note that each domain is a system of behaviors, each requiring multiple tasks to function in the domain. This is proposed as one criterion for defining a domain. They are: mate selection, attachment and caring, pecking order, prey defense, predator action, way finding, food sharing, migration, communication, social cohesion, recognition, food selection, and choice in foraging.

EXAMPLES OF COMPARATIVE COMPLEXITY IN SELECTED DOMAINS AND SPECIES

This brief section has two purposes. The first to indicate how *domains* can be comprised of discrete tasks, and how they can be conceptualized as modules.

The second is to lay groundwork for the construction of indexes described in the proposed theory. To do that, it offers some specific examples to (a) indicate in a concrete way that scoring different animals' acts is possible, and (b) underscore the point that the Model of Hierarchical Complexity applies to universal task complexity, regardless of content and context (for some additional scored animal examples, see Commons, 2006). These indicate why it is useful to distinguish domains within species. By describing tasks that are domain specific, they illustrate that similar task domains exist across species. They also suggest that animals do not vary significantly in their within-species stages of performance.

This set of examples indicates comparisons of some tasks that African Grey parrots, crows, and other birds can perform. Crows perform at the Sentential stage 5. They can string together a sequence of Nominal stage 4 actions, for example, by planning to bend a wire to reach around a corner in a plastic tube to get food. African Grey parrots also operate at stage 5; they can sequence words themselves, and understand word sequences of others. In doing so, it is evident that they can distinguish between the passive voice and active voice. At the Sentential stage, however, these Parrots perform many kinds of tasks that exceed those of crows. One example is saying letters and numbers in order, and counting small sets of objects systematically. Other birds do tasks at earlier stages. At Nominal stage 4, pigeons can switch which key they select and use when a few examples from a class are switched. They construct many arbitrary concepts, including such abstract examples as inside and outside. They are successfully trained to name classes of actions, that is, fish (respond quickly) and non-fish (respond slowly). When a few examples are switched, thus switching the class name from respond quickly to respond slowly, they can switch their rate for all members of the class. Sparrows, by contrast, cannot perform tasks at this stage. Animals' actions have been scored up to the Concrete stage 8. The following are examples of animals' highest stages of performances up to and including that stage.

1. *Sensory or Motor:* When water moves, mollusk opens shell. Reflexively, if something touches membrane, closes shell. Mobile animals (e.g., Aplysia) habituate and sensitize, reflecting classical and operant conditioning. This results in generalizing about which stimuli will elicit the responses of interest
2. *Circular Sensory-motor:* Animals coordinate either perception with an action, or two or more actions, for example, capture and strike movements in the praying mantis (Corrette, 1990). Animals that hunt (e.g., most predatory fish, insects) have behavior controlled by consequences.
3. *Sensory-motor:* Coordinate very basic concepts such as oddity learning in rats (e.g., Bailey and Thomas, 1998), where rats discriminated the "odd" one when given two ping pong balls with food odors and one different order.
4. *Nominal*: Named concepts are used. Vaughan (1988) trained pigeons to associate two arbitrary subclasses of slides of trees with different response rates. High response rate was associated with slides in one subclass and low response rate with slides of the other subclass. When slides in the subclass previously associated with high response rate became associated with low

response rate (and vice versa), the pigeons changed their associations and correctly responded to each slide after a short reacquisition trial, showing they could attach a virtual label to a subclass.

5. *Sentential:* Pepperberg's (1992) African Grey parrot Alex (which died in 2007 at the age of 30) could count two objects ("one, two") and speak in sentences that organized nominal labels and words. When a new question was introduced, "What matter [is this] four corner blue [object made of]?" he correctly responded, "wood." Dogs and cats can perform long arbitrary sequences of actions.
6. *Preoperational:* Rhesus monkeys would be trained to indicate the larger of two sets of 1 to 4 squares and circles in two rows (Brannon and Terrace, 1998). Chimpanzees put nuts onto selected flat anvil stones, and cracked them with selected hammer stones (Inoue-Nakamura and Matsuzawa, 1997). Similar to Hunt's (1996, 2000) observations of similar crow behavior in the wild, New Caledonian crows make tools by bending a straight piece of wire and then use the wire to pull food out of a tube (Weir, Chappell, and Kacelnik, 2002).
7. *Primary:* Rhesus monkeys were trained to select Arabic numerals associated with a number of food pellets (Washburn and Rumbaugh, 1991). They could choose the correct numeral associated with the larger number of food pellets in a random array of up to 5 numerals. Rumbaugh, Hopkins, Washburn, and Savage-Rumbaugh (1989) showed an adult female chimpanzee removing from a TV display the number of boxes appropriate to the value of a randomly selected Arabic numeral, 1, 2, or 3.
8. *Concrete:* Kanzi, a captive Bonobo chimpanzee, used sharp stone flakes and tested the sharpness of each flake with his lips, rejecting non-sharp ones (de Waal and Lanting, 1997). He then made flakes by throwing a rock against a hard surface, producing many flakes at once. Making simple flake tools is a primary order action. Testing the tools is another primary order action. Coordinating one primary stage action with another is a concrete stage action. De Waal (1996) describes how a beta male chimpanzee broke up conflicts in an impartial manner. To act impartially, the beta male had to consider the perspectives of the other chimps along with his own perspective. While his awareness of each of these perspectives is a primary action, his ability to integrate all of these perspectives together demonstrated concrete stage behavior.

Humans can perform such tasks while still children. As human age increases, more possible tasks and domains of activity accumulate. For the measurement theory proposed here to have utility, a taxonomy of domains will be one necessary step. Mascolo ("The Concept of Domain in Developmental Analyses of Hierarchical Complexity," this issue) discusses domains and also the demand to treat domain combinations in certain analyses. King, Kitchener, Wood, and Davison (1990) compared development in domains longitudinally. Another approach to identifying and describing social domains is in Commons (2006), in which size of social grouping was an organizing factor to identify subdomains. Those included self, dyads, triads, committee, small organizations, small markets, organizations, governments and large markets, and societies and cultures. Domains of knowledge provide another angle for conceptualizing domains. For example, physical science, biological science, analytics, experimental, and physical skills. There are a number of issues in identifying and examining domains in the context of humans. See Mascolo (this issue), and Commons (2006) for more discussion.

A CROSS-SPECIES, CROSS-DOMAIN MEASUREMENT THEORY OF COGNITIVE PERFORMANCE

Motivation for the Present Theory

The aforementioned unreconciled arguments are one motivation for the measurement theory presented here. The omission of a systematic connection of the human and animal intelligences in comparative cognition, discussed earlier, is another. There has not been an effective way to compare cognitions of different animals, largely true also for comparing people.

Locurto's (2004) interesting findings were yet another, and led directly to the current proposal. His study of the structure of early acquisition of behavior and of stimulus control has focused on individual differences in mouse "cognition." His earlier findings did not show the kind of robust general factor (i.e., first principal component) typically found in human testing. Instead, he observed a more modular structure. The tasks in those earlier batteries required multiple sessions to complete. The design was changed in light of literature that suggested that clearer evidence of a general factor may be found by running each task for only a few trials, capturing early acquisition performance. The new study had each task designed to provide evidence of learning within a few trials. Because each task was distinct in terms of motivation, sensory modality, and/or behavior measured, the design was robust for testing the presence of a general factor. Results still did not suggest strong support for a general factor. The solution proposed in what follows is a way to conceive of g in animals and people, and a way to measure an approximate g.

PREMISES

This theory relies on the Model of Hierarchical Complexity and its premises. Hierarchical complexity is a measure of one major kind of task difficulty. There are four ways this approach differs from current intelligence measurement procedures. First, hierarchical complexity of tasks forms a content-free, absolute scale rather than one based on norms, context, or content. Second, its formulation is similar to other measures from measurement theory (e.g., Krantz, Luce, Suppes, and Tversky, 1971). Third, it separates the empirical stage of *performance* of tasks from the largely analytic hierarchical complexity of tasks. Finally, it defines developmental stage as performances on tasks accomplished that have a specified hierarchical complexity, instead of defining stage on the basis of some inferred mental or logical operation.

A central premise is that a theory of comparative cognition needs to include multiple indices. The present proposal does so. An index would consist of a numerical scale that is used to compare variables with one another or with some reference number. The Model of Hierarchical Complexity provides such a scale. A theory of comparative cognition also needs to include some measure of general intelligence (g) (Jensen, 1998; Kanazawa, 2004) for animals. The present proposal includes that, and also presents an alternative process and corresponding indexes that may be used along with traditional measures in humans such as IQ (Intelligence Quotient). This approach to human intelligence more closely parallels the assessment of animal intelligence, bridging the current divide.

Another premise is that breadth of intelligence is a better way to summarize the generality of intelligence than the factorial approach. Breadth offers continuity with other animals and incorporates what has become known as the multiple intelligences of Gardner (1983, 1993) and of Sternberg (1985). Breadth, which much of this article leads to defining, will be carefully defined in what follows and distinguished from traditional views of g. A final premise is that a general measurement theory must respect modularity.

INDEXES TO MEASURE AN APPROXIMATE g AND VARIANTS

This section necessarily omits instruction for constructing test items using hierarchical complexity and the role of Rasch Analysis (but see Commons, Rodriguez et al., 2007 and Commons, Goodheart et al., 2007). Starting with the standard tasks within the standard domains, one can construct an analogue of g. There will be three types of measures: (a) the highest stage of performance attained in each domain (HS) including the highest stage in any domain (HHS); (b) a form of g that is somewhat akin to human g; (c) a derived measure of generality of performance, g breadth (gB).

Determining the Rasch Scaled Stage of Performance from the Known Hierarchical Complexity of Stimulus Items

One can construct both an item and a participant stage of performance table from the Rasch scores if the Rasch scores for the task items are reasonable. That means that the order of analytically determined order of hierarchical complexity reasonably predicts the Rasch scaled scores for those items. It also helps if the items are roughly equally spaced but it is not necessary. For example, in the Helper Person Problem, there was a correlation between Rasch scores and Order of Hierarchical Complexity of the items of .998. With such linearity, to construct a table or find a particular participant stage score or item stage score, one runs the regression backward by having Rasch scaled scores predicting Order of Hierarchical Complexity. This yields the regression equations for determining stage for both items and participants. Order of Hierarchical Complexity forms an absolute ordinal scale. It does not require norms. The regression equations have two parameters, a_0 (offset) and a_1 (slope). The predicted Rasch scores are the Rasch stage score that represent the Rasch scores in terms of order of the

Order of Hierarchical Complexity. This yields a stage score instead of raw Rasch score. This generates the table. For any Rasch scaled score for a participant, one has a stage score based on the corresponding order of hierarchical complexity. The stage scores are continuous because the Rasch scales are. This allows one to see transition scores clearly. If the Orders of Hierarchical Complexity are not very evenly spaced, one would linearly interpolate between Orders of Hierarchical Complexity using mean order if there were more than one item per order.

There are two tests of g then. First, the new one is the *r or beta* found by regressing Rasch scaled stage scores obtained for sets of items from multiple domains the items as above. The higher the *r or beta* the more the performances are on the single dimension captured by the Order of Hierarchical Complexity. The second is from a factor analysis of those items. We get much higher $r's$ for the Rasch scaled stage scores than from the raw Rasch scores. This is because effect of the rather arbitrary standard deviations (*SD*) and arbitrary 0 on the Rasch scale is removed and the Rasch scaled stage scores are put on absolute hierarchical complexity scale. There is the problem of some assumption of linearity but also in some sense this is a good test of it.

Constructing a Measure of g

Applying a transform to the Rasch scaled scores for the items and for the participants' performances is useful in comparing performances on difference sequences and in different domains. To form a general measure, first one has to correct for the relatively arbitrary Rasch scale parameters. This is done by translating the Rasch scores into stage scores based on the corresponding absolute values of the order of hierarchical complexity of the items. As mentioned earlier, one finds the regression equation for hierarchical complexity versus Rasch scores. This corrects for the somewhat arbitrary spacing and offset of the Rasch scale. This is also useful in plotting Rasch scaled scores from different instruments versus hierarchical complexity. Without this correction, one can get different zero point because of sample differences and also different *SD*. Even so, this transformation will remove the effect of different *SD* for items/participant or subject.

Highest Stage of Performance Attained in Any Domain

An animal species may be characterized by the highest stage of performance observed with any amount of training on its best task series (HHS).

This first index requires some information as to what the domains are and what the tasks are within each domain. This is currently the area needing development: we know what the tasks accomplish, but we do not have a systematic way to classify domains. Each task has a hierarchical complexity. The highest stage of performance (HS) is just the highest hierarchical complexity of the task that the organism in the species correctly addresses. Then one finds the domain and task in which the highest stage of performance (as determined by hierarchical complexity) occurs (HHS). This will be one number that falls on the stage scale that runs from 0 to 14.

The Index g

The second index, g, is the average of the highest stage numbers of performance in each domain (HS). This is somewhat akin to human g, but g would separate the highest stage from how broad g would be. The average has advantages of the total g, because the average is less sensitive to failing to include a domain or misidentifying a domain. This average of highest stage falls on the stage scale that runs from 0 to 14.

The Index g-Breadth

The third index, called g breadth (gB), measures how broad an organism's capability is by using a scheme that uses a renormed g that removes the effects of the highest stage. This renorming does not refer to a sample but to the process of dividing the average of highest stage in each domain (g) by the top stage of the animal (HHS). This renorming takes away the effect of highest stage. Then we have three numbers, the highest stage (HS); the average stage across domains (g); and g breadth (gB).

Within-Domain Intelligence

The foregoing pertained to measuring an approximate g, by definition a general factor. Another form of intelligence shows up within domains. This within-domain form is like the subtasks within the verbal IQ tasks. The within-domain form shows flexibility of stage of performance (fS) within each domain. One chooses the domain and task in which one wants to measure flexibility, then finds the highest stage of performance (HSdomain) on a wide variety of tasks that occur within that domain (as determined by hierarchical complexity). One then averages the stage numbers of the task performances within the domain. That is g-domain. That is divided by the HSdomain. Again, this scale will consist of the numbers from 0 to 14. That gives gdomain Breadth.

THE HIERARCHICAL COMPLEXITY OF THE CONCEPTS, DEBATES, AND THE PROPOSED MEASUREMENT THEORY TO COORDINATE THEM

The concept *intelligence* originates from Formal stage 10 reasoning. It coordinates the two Abstract stage 9 variables, ability to *know* and to *learn*, both related to the definition of intellect.The more ability one has to know and to learn new things, the more intelligence one is considered to have. Various kinds of knowledge were identified, having different task demands and involving different domains. Factor analysis, mentioned earlier, is a multivariate Systematic stage 11 process to compare relations of such diverse factors. The concept *cognition* may be used by Formal stage 10 as a formal variable to refer to knowledge, the product of learning. This is different from the Systematic stage 11 use of the term, which refers to the *process* of coming to know with multiple variables involved in the process. Metasystematic stage 12 use of the term is evidenced by Maturana and

Varela (1998), for example, who identify and coordinate many systems in their conception of the biological basis of cognition. Such differences in hierarchical complexity play roles in how studies of human and other animal cognition are conceptualized and conducted. Arguments within fields of study can sometimes be found to have their roots in such unrecognized differences as this. When referring to human, animal, or comparative cognition, then, it is important to realize the terms may be interpreted differently at different stages.

The historically predominant separation of studies of human from other animal cognition could have its roots in hierarchical complexity. First, the separation of the study of cognition in humans from other members of the animal kingdom possibly has its roots in Abstract stage 9 ideologies, for example, God's creation of humans as superior to animals (and the rest of creation). Formal stage 10 logics build on Abstract stage variables: if humans are qualitatively superior to animals, then their intelligence or cognition is incomparable. The study of comparative cognition has focused on non-human animals as comparable and excluded humans as incomparable. The context-free task-basis of the Model of Hierarchical Complexity enables the comparison of (even seemingly) disparate systems, such as human and animal cognition. Thus, one role played by the Model in the measurement theory proposed here is a Metasystematic stage 12 function of coordinating systems thus-far perceived as separate.

The arguments in the modularity versus domain-general intelligence debate appear to proceed at last transition step from Systematic stage 11, preceding full Metasystematic stage 12 coordinations of the complexity of organismic action (see "Introduction to the Model of Hierarchical Complexity," this issue). The modularity argument recognizes that one organism has to function in disparate settings and has developed processes that comprise modules. Each module is a system. The modularity argument does not appear to fully coordinate all those systems into an overarching metasystem. The arguments for domain-general, by contrast, have a rationale for an overarching metasystem but have not constructed or identified the disparate systems sufficiently for measurement and accounting for it. Niche-dependent behavior is often not considered. Both sides of the debate have valid arguments and data that support them. The use of Model of Hierarchical Complexity in the measurement theory proposed here enables Metasystematic stage 12 coordination of the domain-specificity of modules and the need for a domain-general measurement mechanism.

CONCLUSION

This theory of g is one demonstration of the Model of Hierarchical Complexity's utility. That utility relies on further research that generates more data than are presently available. To flesh out this theory in detail, such challenges will need to be met. One way to validate such a system of measurement could be to systematically compare a number of individuals (people and animals) using the methods briefly described in this article. However, it takes a great deal of time to test individuals of any species on a large set of tasks. Therefore, it is probably best to analyze the tasks they do and how they do them. By determining the hierarchical complexity of the tasks, one can determine the stage of performance, which requires only scoring. The tasks need to be situated in domains, which need further systematization, particularly for humans. As this develops, it should also inform and perhaps reform some concepts in psychopathology: that many of the problems people suffer from are due to deficits of development in given domains. Such unevenness in development seems to be associated with problems such as criminal activity and substance abuse.

Even with this proposal for a general measurement theory that applies across species and domains, it is only a first step in the process to genuinely address the issues of modularity and generality in intelligence. It is one way to address the controversy, to keep modularity but measure general development across domains. Using such measures will help us understand the evolution of all animals, how they compare, and how and why competencies develop in some domains but not others.

NOTES

1. For example, *Comparative cognition: experimental explorations of animal intelligence*, Eds. Edward A. Wasserman and Thomas R. Zentall. Oxford; New York: Oxford University Press, 2006.
2. Comparative Cognition Society: http://www.pigeon.psy.tufts.edu/ccs/default.htm
3. This article does not address questions or speculations about the existence of cognition or intelligence outside the Kingdom *Animalia.*

REFERENCES

Bailey, A. M., and Thomas, R. K. 1998. An investigation of oddity concept learning by rats. *Psychological Record* 48: 333–344.

Brannon, E. M., and Terrace, H. S. 1998. Ordering of the numerosities 1 to 9 by monkeys. *Science* 282: 746–749.

Carroll, J. B. 1993. *Human cognitive abilities*. New York: Cambridge University Press.

Chiappe, D., and MacDonald, K. 2005. The evolution of domain-general mechanisms in intelligence and learning. *The Journal of General Psychology* 132(1): 5–40.

Commons, M. L. 2006. Measuring an approximate g in animals and people. *Integral Review* 3: 82–99.

———, Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., and Adams, K. M. 2007. Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. In *Rasch measurement: Advanced and specialized applications*, Eds. Smith, E. V. Jr., and Smith, R. M., 121–147. Maple Grove, MN: JAM Press.

———, Rodriguez, J. A., Miller, P. M., Ross, S. N., LoCicero, A., Goodheart, E. A., and Danaher-Gilpin, S. 2007. *Applying the model of hierarchical complexity.* Unpublished manuscript, Dare Institute, available from commons@tiac.net

Corrette, B. J. 1990. Prey capture in the praying mantis tenodera-aridifolia-sinensis-coordination of the capture sequence and strike movements. *Journal of Experimental Biology* 148: 34.

Cosmides, L., and Tooby, J. 1994. Origins of domain specificity: The evolution of functional organization. In *Mapping the mind: Domain specificity in cognition and culture*, Eds. A. Hirschfeld, L., and Gelman, S. A., 85–116. New York: Cambridge University Press.

———, and Tooby, J. 2007. *Evolutionary psychology: A primer by Leda Cosmides and John Tooby*, http://www.psych.ucsb.edu/research/cep/primer.html (Accessed 21 September 2007.)

Darwin, C. 1871. *The descent of man, and selection in relation to sex*. 2 vols. London: John Murray. http://human-nature.com/darwin/descent/contents.htm (Accessed 20 August 2006.)

de Waal, F. B. M. 1996. *Good natured: The origins of right and wrong in humans and other animals*. Cambridge: Harvard University Press.

———, and Lanting, F. 1997. *Bonobo: The forgotten ape.* Berkeley: University of California Press.

DiLalla, L. F. 2000. Development of intelligence: Current research. *Journal of Psychology* 38: 3–7.

Galton, F. 1962. *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan/Fontana. (Originally published 1869, 1892)

Gardner, H. 1993. *Frames of mind: The theory of multiple intelligences.* New York: Basic Books. (Originally published 1983)

Geary, D. C. 2004. *Origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: APA Books.

Gintis, H. 2007. Book Review: David J. Buller. 2005. Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature. MIT Press, Cambridge, MA. *Journal of Bioeconomics* 9: 191–199.

Hunt, G. R. 1996. Manufacture and use of hook-tools by New Caledonian crows. *Nature* 379: 249–251.

———. 2000. Human-like, population-level specialization in the manufacture of pandanus tools by New Caledonian crows Corvus moneduloides. *Proceedings of the Royal Society of London B* 267: 403–413.

Inoue-Nakamura, N., and Matsuzawa, T. 1997. Development of stone tool use by wild chimpanzees (pan Troglodytes). *Journal of Comparative Psychology* 111(2): 159–173.

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

Kanazawa, S. 2004. General intelligence as a domain-specific adaptation. *Psychological Review* 111(2): 512–523.

King, P. M., Kitchener, K. S., Wood, P. K., and Davison, M. L. 1990. Relationships across developmental domains: A longitudinal study of intellectual, moral, and ego development. In *Adult Development: Vol. 1: Comparisons and applications of developmental models*, Eds. Commons, M. L., Sinnott, J. D., Richards, F., and Armon, C., 57–72. New York: Praeger.

Krantz, D. H., Luce, R. D., Suppes, P., and Tversky, A. 1971. *Foundations of measurement: Vol. 1. Additive and polynomial representations*. New York: Academic Press.

Locurto, C. 2004. The structure of early acquisition. Paper presented at the Comparative Cognition Meeting, March 22–25, 2006, Melbourne, FL.

Maturana, H. R., and Varela, F. J. 1998. *The tree of knowledge: The biological roots of human understanding.* Boston: Shambhala.

Miller, G. F. 2000. Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. In *Evolutionary approaches to human reproductive behavior*. *Annals of the New York Academy of Sciences*, vol. 907, Eds. LeCroy, D., and Moller, P., 62–74. New York: New York Academy of Sciences.

Pepperberg, I. 1992. Proficient performance of a conjunctive, recursive task by an African gray parrot (Psittacus erithacus). *Journal of Comparative Psychology* 106(3): 295–305.

Romanes, G. J. 1888. *Mental evolution in man: Origin of human faculty*. London: Kegan Paul, Trench.

Rumbaugh, D. M., Hopkins, W. D., Washburn, D. A., and Savage-Rumbaugh, E. S. 1989. Lana chimpanzee learns to count by "NUMATH": A summary of a videotaped experimental report. *Psychological Record* 39(4): 459–470.

Shalizi, C. R. 2007. *g, a statistical myth.* http://www.cscs.umich.edu/~crshalizi/weblog/ (Accessed 21 October 2007.)

Skottke, K. R. 2006. *The evolution of human intelligence: Increasing importance of domain-specific intelligence in the modern environment.* http://www.personalityresearch.org/papers/skottke.html (Accessed 22 May 2006.)

Spearman, C. 1904. "General intelligence," objectively determined and measured. *American Journal of Psychology* 15: 201–293.

Sperber, D. 2002. In defense of massive modularity. In *Language, brain and cognitive development: Essays in honor of Jacques Mehler*, Ed. Dupoux, E., 47–57. Cambridge: MIT Press.

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

Thagard, P. 2006. *Hot thought: Mechanisms and applications of emotional cognition.* Cambridge: MIT Press.

Thorndike, E. L. 1898. Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review,* Monograph Supplements, No. 8. New York: Macmillan.

———. 1904. *Introduction to the theory of mental and social measurements*. New York: Teacher's College.

———. 1911. *Animal intelligence*. New York: Macmillan.

———. 1920. Intelligence and its use. *Harper's Magazine* 140: 227–235.

———. 1927. *The measurement of intelligence*. New York: Teachers College Press.

Thurstone, L. L. 1931. Measurement of social attitudes. *Journal of Abnormal and Social Psychology* 26: 249–269.

Uttal, William R. 2003. *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge: MIT Press.

Vaughan, W. 1988 Formation of equivalence sets in pigeons. *Journal of Experimental Psychology—Animal Behavior Processes* 14(1): 36–42.

Washburn, D. A., and Rumbaugh, D. M. 1991. Ordinal judgments of numerical symbols by macaques (Macaca mulatta). *Psychological Science* 2(3): 190–193.

Wechsler, D. 1967. *Manual for preschool and primary scale of intelligence.* New York: The Psychological Corporation.

Weir, A. A. S., Chappell, J., and Kacelnik, A. 2002. Shaping of hooks in New Caledonian crows. *Science* 297: 981.